

基礎統計学Ⅲ

自然科学の統計学

東京大学教養学部統計学教室 編

慶應義塾大学環境情報学部
富田研究室
〒252 藤沢市遠藤5322
TEL 0466-47-5111

東京大学出版会

「基礎統計学」の刊行にあたって

【執筆分担】

- 第1章 矢島美寛
- 第2章 廣津千尋
- 第3章 藤野和建
- 第4章 廣津千尋
- 第5章 廣津千尋
- 第6章 竹村彰通
- 第7章 竹内 啓, 藤野和建
- 第8章 繩田和満
- 第9章 松原 望
- 第10章 矢島美寛
- 第11章 伏見正則

Analysis of Scientific Data
(Basic Statistics 3)
Statistics Section, Department of Social Sciences, College
of Arts and Sciences, University of Tokyo

University of Tokyo Press, 1992
ISBN978-4-13-042067-9

統計学を学ぶ目的は、統計的なものの考え方や統計手法の基礎を学ぶことである。その内容は、記述統計学、確率・確率分布と統計的推測の基礎的な部分、統計調査と調査データについての部分、さらに数理統計の基礎の部分とに分けることができる。もちろん、それぞれの部分については、学ぶ人の基礎知識、ことに数学に関する基礎知識に応じて、さまざまなレベルがあるし、また各応用分野において用いられるデータおよび統計手法も異なるのが、ふつうである。

このような統計学を教える立場にある者からすると、以下のような点が問題である。

1. 各専門分野さらには実社会において、統計的なものの見方や統計手法の必要性が高まっているにもかかわらず、統計学を敬遠する人が少なくない。ことに文科系の学生には、統計学は数学の一部という先入観からか、その傾向が強い。

2. 一方で、統計学の内容を、数式を省略してその結果のみを覚えようとすると、確率や統計の基本的な考え方を理解するのが難しくなるし、他方、あまり数理的な面だけに注目しすぎると、応用数学の一部になってしまい、統計調査やその結果としてのデータの問題など、現実との接点を見失ってしまうこともある。

3. コンピューターのハードウェア、ソフトウェアの急速な進歩に従って、統計学、とくにそのうちのデータ解析と情報処理が融合しつつあり、統計学の学習とコンピューターの操作が区別できなくなるおそれが出てきている。

これらの点は、根本的には高校教育や大学入試での統計学の扱われ方にかかりがある。それはそれとして、なるべく多くの人が統計学に興味をもつよう、教科書や講義の内容をよくする努力がなされるべきである。その際、統計学への親しみと論理性、そしてその数学的な厳密性と現実への適用などのバラ

ンスをどのようにとるかということが、もっとも重要な問題となる。

第3の点も、統計学の講義を、黒板とチョークのみを用いたものからコンピューターとくにパソコンを用いた、より視覚的かつ実践的なものに作り変えていかなければならない。同時に、データさえ入力すれば自動的に結果が出てくるという環境のもとでは、コンピューターが計算した結果をうのみにするのではなくて、それを理解し評価する能力を養うことが絶対的に必要な課題となる。

このような問題意識から、われわれは、統計学の基礎に関する、スタンダードな（もちろん、あくまでその一つの）教科書として「基礎統計学」（全3冊）を刊行することにした。このシリーズの基本的な目的は、上に述べた問題点の解決にとり組むことである。それをこのような形で公けにすることにした理由は、これらの問題は、必ずしも、特定の大学の特定の学部における統計学の講義だけの問題ではなくて、統計学を学んだり、教えたりする多くの人々に共通の問題であると考えたからである。

「基礎統計学」の基本方針とその構成は、以下のとおりである。

i. 「基礎統計学」を I 「統計学入門」、II 「人文・社会科学の統計学」、III 「自然科学の統計学」の3冊に分ける。

ii. I 「統計学入門」は、文科系・理科系両方の学生を対象として、統計学の初心者に統計的なものの考え方をやさしく解説し、統計学に関する体系的な知識を与えることを目的としている。そして、その内容は、統計学の基礎概念、データの整理（記述統計学）、確率・確率分布の基礎、母集団と標本の考え方、標本分布、推定と検定の基礎（統計的推測）、回帰分析の初步である。

iii. II 「人文・社会科学の統計学」は、主として、文科系の学生を対象として、人文・社会科学で用いられる統計手法を、データの取り方と分析を中心とし、やさしく解説している。その内容は、データ論、標本調査、主な経済データとその分析、主な社会・心理データとその分析である。

iv. III 「自然科学の統計学」は、主として、理科系の学生を対象として、自然科学・工学・医学等への応用を念頭におき、さまざまな統計的な考え方を紹介し、その基礎をわかりやすく解説している。その内容は、線形回帰モデル、実験データの分析、推定と検定の理論、ノンパラメトリックな手法、質的データの分析、ベイズ決定、確率過程論である。

v. 上記3冊の内容については、統計学に対する親しみを増すとともに、それに対する興味がわくように、その書き方にも注意を払っている。また統計学を履修しない学生の自習用テキストにもなるよう配慮している。

vi. 上記3冊の分量は、いずれも、ほぼ半年（約15週間）分であって、大学の単位数でいうと、2単位に相当する量に相当している。

この基本方針と構成は、主として、上記の問題点1と2に対応するためのものであって、これだけでは、問題点の3に関しては十分に対応しきっていない。今後この「基礎統計学」に即した統計ソフトウェアの開発、および演習用・実習用の小型のデータベースの構築などをめざす予定である。さらに、このシリーズ全体を通じた統計学小辞典ないしはハンドブックも、準備する必要がある。

この「基礎統計学」に関しては、企画の段階から、東京大学の各学部で統計学の講義を担当しておられる先生方にお世話になった。このシリーズは、形のうえでは、東京大学教養学部社会学科統計学教室編になっているが、これらの先生方のご協力によるところが大きい。とくに、直接ご執筆いただいた先生方、序文をお書きいただいた竹内啓先生、この企画の発足時より、積極的な推進役であり、現在は先端科学技術研究センターに移られた廣松毅先生をはじめ東京大学統計学連絡委員会の委員の先生方、そして、われわれの構想に関するアンケートに懇切丁寧にお答えいただき、かつ建設的なご意見やアドバイスをいただいた先生方に心から御礼申し上げたい。同時に、当然のことながら、われわれは、今後も東京大学の一般教育課程における統計教育の発展に力を注ぐつもりである。さらなるご協力を願う次第である。

1991年5月

東京大学教養学部社会学科統計学教室
高橋伸夫、中井檢裕、繩田和満、藤野和建、松原 望
(1991年4月1日現在)

序 文

統計学とは何かということについては、昔からいろいろ面倒な議論がある。簡単にいえば、それは数字データというものを、どのように分析し、どのような判断をくだしたらよいかを論ずる学問であるといってよい。

数字データとは、何らかの客観的な対象の性質、あるいは現象を数字で表したものである。それは或る「もの」の個数、長さ、体積、重さ、あるいは或る「こと」の起った回数、時刻、続いた時間の長さなどを表したものである。

これらは、純粋な数学的な値、たとえば円周率 $\pi=3.1415926535\cdots$ 、あるいは $e=2.718281828\cdots$ などとは異なった性質をもっている。このような値は厳密に数学的に定義され、少なくとも原理的には数学的に何ヶタでも計算することができる。「円周率」は「円の周囲の長さと直径との比」と定義されても、それは現実に描かれた円の周囲と直径とを測定して得られるものではなく、理論上の「理想的な円」についての値である。

これに対して数字データは、つねに何らかの具体的な「もの」や「現象」について、測ったり数えたりしてはじめて得られるものである。したがって、それは対象の具体的な性質も反映しており、たとえば長さ、重さ等の単位をもつものとして表される。

同時にそれはまた、何らかの観測、実験、調査を通じて得られるものであるから、それがどのようにして行われたかということも数字データに反映する。

そこで数字データというものは、数学的な値のようにいくらでも厳密に求めることはできない。長さや重さの測定とか、実験結果の観測にはつねに誤差というものがともなう。また計器の目盛り等の細かさ等にも限界があるから、現実に求められるヶタ数はそれほど大きくはできない。

また仮に測定がいくらでも精密にできるとしても、実際にむやみにヶタ数を多くしても無意味であることが少くない。たとえば、ある人の体重を測定する

序 文

とき、グラムの単位まで測っても無駄であろう。人の体重は、飲食や排泄、あるいは発汗などによって数十～数百グラムは絶えず変動しているからである。あるいは、東京から大阪までの距離をメートル単位まで測ってもしかたがない。そこまで定めようとすれば、東京と大阪のどの地点からどの地点まで、どのようなコースをたどって測った距離であるかを厳密に決めなければならないが、仮にそんなことができたとしても、それは馬鹿げたことではなかろうか。

実際われわれが数字データから知りたいのは、そんなに厳密な値ではない。体重については、0.1 kg の単位まで知れば十分だし、東京大阪間の距離は、たとえば東京駅から新大阪駅まで、新幹線で何 km ということがわかれればよいのがふつうである。

もちろん物理的な定数、たとえば真空中での光の速さのように、理論上つねに厳密に一定であると考えられ、したがって原理的にいくらでも精密に測定できると思われる量もあるが、数字データの表すものとしては、むしろそれは例外であって、あまり厳密に求めることはできないし、またその必要もない場合が少くない。

また、われわれが知りたいことが、実際そこにあるものの性質、あるいは現に起った現象そのものではなくて、その背後にあるより一般的な事物の性質や傾向であることも少くない。たとえば、ある工業製品の性質を調べるとき、知りたいのは実際にテストされたり観測された標本の性質ではなく、その標本と同じ、あるいは同じとみなされる製品の性質である。ところが、そのような製品の性質は必ずしも一定ではないかも知れない。そうすると、知りたいのは、そのような製品の多数についての平均の性質、あるいはその中で一定の条件をみたすものの比率である。ところが標本として観測された製品の性質は平均的と必ずしも一致しないであろう。そこで何個かの標本の性質から、多くの製品の平均的な性質についての判断をくださなければならぬことになる。

また、最初からわれわれの関心が個々の対象の性質ではなく、同じような対象の集団全体の傾向である場合もある。たとえば、小学生の体重を測定した場合、1人1人の児童の体重よりも、あるいはそれ以外に、日本のある学年の児童全体の集団について、その体重の分布がどうなっているかを知りたいことがあるかも知れない。このような場合には、多数の児童の体重測定の結果をまとめるかも知れない。

序 文

めて、その全体の傾向を知るために、平均や標準偏差を計算したり、あるいはヒストグラムを描いたりすることが必要になる。このような場合には1人1人の児童についての測定値は、それが集団全体の様子を何らかの意味で反映しているという点からのみ扱われることになる。

そこで数字データといふものには、つねに何らかの意味で、対象についてわれわれが知りたいと思う性質とは無関係な変動が加わっていると考えねばならない。このような変動は、誤差、ばらつき、ノイズなどといふ言葉で表現されるのがふつうである。

ところで、このような変動は偶然的なものと見なされることが多い。そして更に偶然的な変動は、確率論の法則に従うものと考えられるのが普通である。

数字データを確率的な偶然変動をふくむものと見なすと、それをどのようにして分析し、それからわれわれが知りたいと思う事実について、どのようにして判断をくだしたらよいかを数学的に明らかにすることができる。このような数字データを扱うための数学的理論を与えるのが、数理統計学である。

数理統計学は数学的理論ではあるが、現実の数字データを分析するには、ただその数学的方法を形式的にあてはめればよいというものではない。具体的な対象の性質と分析の目的に応じて適切な方法を選び、その結果を正しく判断しなければならない。そういう意味では応用する分野ごと、あるいは問題ごとに数理統計学の方法も異ならなければならないと思われるかも知れない。実際、いろいろな分野には、それぞれ特有の方法や理論をもつ数理統計学の体系がつくられている。それは計量〇〇学といふ言葉で表されることが多い。たとえば、計量生物学、計量経済学、計量心理学等々である。

しかし、これらのすべての分野を通じて、またすべての数字データの解析に対して、共通に適用できるような数学的方法が、最近200年くらいの間につくられてきた。それが数理統計学の一般理論あるいは基礎理論である。このシリーズは、その概略をこの分野での予備知識を前提にせずに解説したものである。

現在では、自然、社会、人文諸科学、あるいは工学、医学等はほとんどすべての研究分野で数字データが用いられており、その意味では、数理統計学の最小限の知識は、ほとんどすべての分野で必要とされているといってよい。このシリーズは、その意味ですべての分野の学生、研究者に数理統計学の基本的な

知識を与えることを目的としている。

現在では、数理統計学の方法を応用した場合の計算はほとんどすべてコンピューターで容易に行なうことができる。また、そのためのプログラムパッケージもいくつもつくられている。場合によると、方法の意味を理解しないでも、コンピューターが結果を出してくれることもある。しかし、方法の意味がわからずには、ただ計算をしてもその結果を正しく利用することはできないし、また場合によっては誤った判断をくだすことになる危険がある。数理統計学の方法を形式的に説明するだけでなく、その意味を把握すること、それがこの教科書の最も重要な目標である。このシリーズの著者たちはそのために工夫をこらしたはずである。

多くの分野の学生諸君や、あるいはいろいろな学部で統計学を講義する教授諸氏にとって、このシリーズが有益な教科書として役立つことができれば幸いである。

1991年5月

竹 内 啓

本書の使い方

本書は『基礎統計学』シリーズの第III巻であって、主として自然科学、工学、医学などの分野で用いられている統計的方法を、比較的新しいものも含めて、一般的に解説したものである。この意味で、本書の主たる対象は理科生である。しかし、ここで取り上げた方法の多くは、人文・社会科学の分野においてもよく用いられており、本書はそうした方法に興味をもつ文科生にとっても役立つであろう。

本書の各章は独立しているので、読者は自分の関心に応じて適当な章から始めればよく、第1章から順番に読んでいく必要はまったくない。ただし、本書の内容およびレベルは、本シリーズの第I巻『統計学入門』程度の予備知識を前提としているので、それぞれの章を開く前に本シリーズの第I巻ないしはそれと類似の確率・統計に関する標準的な教科書に一通り目を通しておくことを、お勧めしたい。

参考までに、考えられそうな関心領域に応じた、本書の使い方を例示しておこう。

1. データの分析に关心がある場合：第3, 4(とくに4.1, 4.2節), 5, 7, 8章。
2. 回帰分析ないしは線形モデルに关心がある場合：第2, 4, 8章。
3. 推定、検定に关心がある場合：第4, 5, 6, 7章。
4. 確率、確率過程、OR、シミュレーションに关心がある場合：第1, 10, 11章。
5. ベイズ統計学ないしは意思決定理論に关心がある場合：第9章。

その他、読者個人の関心に応じて、さまざまな使い方があるであろう。

各章にある小字の部分(理論的関心に対する部分)、編集者が挿入した囲み記事(歴史的エピソードや日常的な例など)あるいは練習問題、細項目を付した索引などの目的や扱いは、本シリーズの他の巻と同様である。学習の目的や動機、予備知識に応じて、それらを有効に役立てていただければ、幸いである。

なお、練習問題は本文の理解の補助というだけでなく、パソコンによるデータ分析の実例を提供するという意味も含めた。計算の実行はパソコンを用いることを前提としているので、その習熟のためにも大いに活用されたい。

†印のついた表、図、練習問題などは、筆者の一人の次のウェブサイトに登載されている(あるいは登載予定の)ものである。

参照サイト <http://www.qmss.jp/qmss/related/stat03.htm>

記号表

α	第一種の過誤の確率, 有意水準	$N(0, 1)$	標準正規分布
$Be(\alpha, \beta)$	ベータ分布	n	標本の大きさ
$Bi(n, p)$	二項分布	Π	射影子
$B(\alpha, \beta)$	ベータ関数	$P(\cdot)$	確率
b_i, β_i	回帰係数	$Po(\lambda)$	ポアソン分布
b_1, b_2	標本歪度, 標本尖度	nPr	順列の数
β	第二種の過誤の確率	p, q	確率の値 ($q = 1 - p$)
$\beta_s(\cdot)$	検出力	r, r_{xy}	標本相関係数
β_1, β_2	母歪度, 母尖度	R	重相関係数
nC_r	組み合わせの数	rank	行列のランク, 階数
$Cov(\cdot, \cdot)$	共分散(確率変数の)	r_s, r_K	順位相関係数
$D(\cdot)$	標準偏差(確率変数の)	r_h	自己相関係数
$E(\cdot)$	期待値	ρ, ρ_{xy}	母相関係数
$Ex(\lambda)$	指数分布	S^2	(不偏でない) 標本分散
F	F 分布に従う量, F 統計量	s^2	標本(不偏)分散
$F(n, m)$	自由度 n, m の F 分布	σ^2	母分散
$F_a(n, m)$	同パーセント点	t	t 分布に従う量, t 統計量
$F(x)$	累積分布関数	$t(n)$	自由度 n の t 分布
f_i, f_{ij}	度数	$t_a(n)$	同パーセント点(上側)
$f(x)$	(確率) 密度関数	tr	行列のトレース
$\phi(z)$	標準正規分布の累積分布関数	θ, Θ	母数
$Ga(\alpha, \lambda)$	ガンマ分布	θ (添字)	θ に対する P, E, V, f, I の指定
$\Gamma(s)$	ガンマ関数	$V(\cdot)$	分散
H	仮説	x, x_i, y, z	変数, 変量, 実現値
$I(\theta)$	フィッシャー情報量, 情報行列	\bar{X}, \bar{x}	標本平均
η^2	決定係数	X, Y, \dots	確率変数
χ^2	χ^2 分布に従う量, χ^2 統計量	X	計画行列
$\chi^2(n)$	自由度 n の χ^2 分布	z	偏差値(得点)
$\chi^2_a(n)$	同パーセント点	Z	標準正規分布に従う変数
$L(\theta)$	尤度関数	ω	根元事象
$M_x(t)$	モーメント母関数	Ω	標本空間
μ	母平均, 期待値(平均)	符号	
μ_r, μ_r'	モーメント(積率)		: 条件(付)
$N(\mu, \sigma^2)$	正規分布	$\hat{\sim}$	推定量, 推定値
		[,]	: (信頼) 区間
		.	: ドット記法

目次

「基礎統計学」の刊行にあたって (i)

序文 (v)

本書の使い方 (ix)

記号表 (xvi)

第1章 確率の基礎

1.1	いろいろな確率分布	1
1.2	モーメント母関数の性質と応用	19
1.3	中心極限定理	20
	練習問題	22

第2章 線形モデルと最小二乗法

2.1	線形モデル	25
2.2	最小二乗法	36
2.3	最小二乗推定量の分散	49
2.4	誤差分散 σ^2 の推定	51
2.5	正規線形モデルと標本分布	55
2.6	線形仮説の検定	59
	[付節] 線形代数と統計学	67
	練習問題	74

第3章 実験データの分析

3.1	2 標本問題	79
3.2	1 元配置分散分析	84
3.3	交互作用と要因実験	94

3.4 2元配置分散分析	96
練習問題	108
第4章 最尤法	
4.1 一般線形モデル	111
4.2 最尤法	115
4.3 データのもつ情報量	120
4.4 最尤推定量の最適性	128
4.5 検定の漸近論	135
練習問題	142
第5章 適合度検定	
5.1 χ^2 適合度検定	145
5.2 2つの二項分布の比較	152
5.3 多項分布の一様性検定	156
5.4 分割表の対称性の検定	162
5.5 ブラッドリー・テリーのモデル	166
5.6 3次元分割表と対数線形モデル	170
練習問題	173
第6章 検定と標本の大きさ	
6.1 検定の検出力	177
6.2 主な検定と検出力	184
6.3 標本の大きさ n の決定	193
6.4 最強力検定	195
練習問題	199
第7章 分布の仮定	
7.1 正規分布の仮定	201

7.2 点推定	206
7.3 仮説検定	213
7.4 正規分布の仮定のチェック	222
練習問題	228
第8章 質的データの統計的分析	
8.1 二値データ	231
8.2 ロジット・モデルとプロビット・モデル	236
8.3 確率の推定	241
8.4 説明変数が2個以上の場合	245
練習問題	249
第9章 ベイズ決定	
9.1 ベイズの定理	251
9.2 事前確率分布と事後確率分布	253
9.3 ベイズ推定	258
9.4 統計的決定理論	264
9.5 ベイズ判別	268
練習問題	273
第10章 確率過程の基礎	
10.1 ランダム・ウォークと破産問題	277
10.2 ブラウン運動	283
10.3 マルコフ連鎖	286
10.4 ポアソン過程と出生死滅過程	293
10.5 確率過程の応用例	298
[付節] 差分方程式の解法	304
練習問題	305

第 11 章 亂数の性質	307
11.1 亂数の性質	307
11.2 一様乱数の発生法	311
11.3 正規乱数の発生法	319
11.4 一般の乱数の発生法	321
[付節] 多次元疎結晶構造とスペクトル検定	328
練習問題	331
統計数値表	333
練習問題の解答	347
参考文献	354
索引	357

基礎統計学 I 統計学入門 目次

第 1 章 統計学の基礎
第 2 章 1 次元のデータ
第 3 章 2 次元のデータ
第 4 章 確率
第 5 章 確率変数
第 6 章 確率分布
第 7 章 多次元の確率分布
第 8 章 大数の法則と中心極限定理
第 9 章 標本分布
第 10 章 正規分布からの標本
第 11 章 推定
第 12 章 仮説検定
第 13 章 回帰分析

第 1 章 確率の基礎

データをもとに、科学的・客観的・効率的な方法で全体の構造を明らかにするのが「統計学」の役割である。しかし同じ条件・環境のもとで実験・調査・観察を繰り返しても常に同じデータを得るとはかぎらない。さまざまな不確実な要因が作用するからである。「不確実性」を支配する確率法則を理解することが、統計学の第一歩である。

1.1 いろいろな確率分布

1.1.1 確率変数と確率分布

実験・調査・観察の結果得られる数量データは、未知の要因・偶然・測定誤差などに影響されてさまざまな値をとる場合が多い。小学校で児童の体力測定を実施すれば、選ばれた児童により当然身長、体重は異なる。できる限り制御された実験室で観測されたデータも、すべてが同じということは不可能に近い。さいころ、ルーレットなどの遊戯でもでる目は明らかに一定ではない。ある変数のとる値が事前に予知できず、確率 P にしたがって出現するとみなすとき、これを確率変数 random variable という。一般に大文字の X, Y, Z などで表現する。

確率変数には、さいころのように、とる値の個数が有限個、あるいは無限個でも、 $0, 1, 2, \dots$ 、と数え上げができるものと、体重や身長のように、ある定められた範囲の任意の実数値をとることが可能なものがある。前者を離散型 discrete type、後者を連続型 continuous type という。

確率分布 probability distribution は、ある確率変数が、どんな値をとる可

表現能性が大きいか、あるいは小さいかを、実現値と確率の間の関数として、表現したものである。実現値は通常小文字 x, y, z などで表現する。とくにある値 x 以下の値をとる確率を知りたいとき

$$(1.1) \quad F(x) = P(X \leq x)$$

と定義する。先ほどの体力測定の例で、 X が身長を表すとすれば、身長が x 以下の中から選ばれる確率を意味する。この確率は、 x に依存してきまる関数であり、累積分布関数 cumulative distribution function あるいは単に分布関数 distribution function とよぶ。 $F(x)$ は x の単調非減少関数である。また、 X は必ず有限の値をとると仮定し、

$$F(\infty) = 1, \quad F(-\infty) = 0$$

とする。 X が連続型の確率変数の場合、確率分布 $F(x)$ は x に関して微分可能な場合が多い。このとき導関数を確率密度関数 probability density function といい、 $f(x)$ と表す。すべての x に対して $f(x) \geq 0$ であり、 $F(x)$ は

$$(1.2) \quad F(x) = \int_{-\infty}^x f(u) du$$

と表現できる。とくに

$$F(\infty) = \int_{-\infty}^{\infty} f(x) dx = 1$$

である。また区間 $(a, b]$ の値をとる確率は

$$(1.3) \quad P(a < X \leq b) = \int_a^b f(u) du$$

となる。ただし、左辺において $<$ と \leq の違いは無視しても差し支えない。なぜなら一点の積分は 0 であるから、すべての x について

$$P(X=x)=0$$

が成立する。ただ離散的な確率変数の場合には、とりうる値の個々の確率で表した方が便利である。いま $x_i, i=0, 1, 2, 3, \dots$ をとりうる値としたとき

$$(1.4) \quad p_i = P(X=x_i), \quad i=0, 1, 2, \dots$$

で定義される関数を、確率関数 probability function という。このとき

$$F(x) = \sum_{x_i \leq x} p_i$$

となるから、 $F(x)$ は、点 x_i において p_i の高さだけジャンプする階段関数である。

ところで確率分布について、分布の中心、あるいはばらつきなどが一目でわかる指標があると便利である。中心の指標のなかで代表的なものは期待値 expectation である。離散的な確率変数に対して、期待値 $E(X)$ は、とりうる値にその確率をかけた加重和

$$(1.5) \quad E(X) = \sum_i p_i x_i$$

で定義する。一方、ばらつきの指標としては、分散 variance が頻繁に用いられる。いま $\mu = E(X)$ とおいたとき、分散 $V(X)$ は

$$(1.6) \quad V(X) = E\{(X - \mu)^2\} = \sum_i p_i (x_i - \mu)^2$$

で定義する。 $V(X)$ は

$$V(X) = E(X^2) - \{E(X)\}^2$$

とも表現できる。連続型の確率変数にたいしては、期待値、分散とともに和が、積分に変わるだけである。すなわち

$$(1.7) \quad E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

$$(1.8) \quad V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

となる。最後に確率変数 $X_i (i=1, 2, \dots, n)$ の線形和

$$c_1 X_1 + c_2 X_2 + \dots + c_n X_n$$

の期待値と分散に対する公式を与えよう。ここで c_1, c_2, \dots, c_n は定数とする。まず、期待値については

$$(1.9) \quad E(c_1 X_1 + c_2 X_2 + \dots + c_n X_n) = c_1 E(X_1) + c_2 E(X_2) + \dots + c_n E(X_n)$$

が成立する。一方、分散については

$$(1.10) \quad V(c_1 X_1 + c_2 X_2 + \dots + c_n X_n) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \text{Cov}(X_i, X_j)$$

が成立する。ここで $\text{Cov}(X_i, X_j)$ は、共分散 covariance とよばれる量で

$$\text{Cov}(X_i, X_j) = E\{(X_i - E(X_i))(X_j - E(X_j))\}$$

で定義する。 $X_i (i=1, 2, \dots, n)$ が互いに独立のときは

$$\text{Cov}(X_i, X_j) = E(X_i - E(X_i))E(X_j - E(X_j)) = 0 \quad (i \neq j)$$

であるから、(1.10) は

(1.11) $V(c_1X_1 + c_2X_2 + \dots + c_nX_n) = c_1^2 V(X_1) + c_2^2 V(X_2) + \dots + c_n^2 V(X_n)$
となる。

1.1.2 離散分布

目的や用途に応じてさまざまな分布が用いられるが、この項においては、代表的な離散分布として、二項分布、幾何分布、負の二項分布、超幾何分布、ポアソン分布を取り上げよう。なお、これらの期待値、分散については、2節でとりあつかっている。

二項分布 データを収集するために行う実験、調査、観測をひとまとめにして試行 trial という。いま、ある試行を繰り返し行っているとしよう。またその結果、特定の事象 S が起きるか否かに関心があるとする。たとえば工場の生産工程であれば、完成品を次々に取り上げ良品か不良品か点検する場面を想像してみよう。このとき $S = \text{「完成品が良品である」}$ となる。S が起きることを「成功」とよび、一方起きないことを「失敗」とよび F と表す。

さて、試行を n 回繰り返したときの成功の総回数を確率変数 X で表す。ここで X の確率分布をもとめよう。さまざまな条件が一定に保たれた環境における、実験・観測を考える場合には、以下の仮定をおくことは自然であろう。まず各回の試行結果は、他の回の試行結果に影響を及ぼさないとする。これは独立性とよばれている。次に成功の確率、失敗の確率が、各試行を通じて一定とする。この性質を定常性という。さらに、ある事象 S が起きるか否かのみに関心がある。これを二値性という。以上、独立性・定常性・二値性の三条件を満足する試行系列をベルヌーイ試行 Bernoulli trial とよぶ。

各回の試行で成功する確率を p 、失敗する確率を q とする。定常性の仮定から、 p, q は一定であり、また確率であることから $p \geq 0, q \geq 0$ で、かつ

$$(1.12) \quad p+q=1$$

をみたす。一般の n に対する確率分布をもとめる前に、見通しをつけるため $n=3$ の場合を考えよう。このとき起こりうる系列は以下の $2^3=8$ 通りある。

SSS SSF SFS FSS SFF FSF FFS FFF
すべて成功、あるいはすべて失敗はひとつしかない。したがって独立性の仮定から

$$P(X=0)=q^3, \quad P(X=3)=p^3$$

である。1回成功し、2回失敗する系列は3つある。各々の系列が起きる確率は、 p と q の順序を入れ換えれば、すべて pq^2 となる。したがって

$$P(X=1)=3pq^2$$

である。同様に

$$P(X=2)=3p^2q$$

となる。

一般の n 回の試行のうち、 k 回 ($k=0, 1, 2, \dots, n$) 成功する確率を求めよう。起こりうる実現値の系列は



という形式の、S と F からなる文字列であるから、総数は 2^n 通りである。そのうち、 k 回成功するする系列は、 k 個の S と、 $(n-k)$ 個の F から成る文字列であるから、組合せの数 ${}_n C_k = n! / \{k!(n-k)!\}$ 個ある。各々の系列が起きる確率は、 p と q の順序を入れ換えればすべて $p^k q^{n-k}$ となる。したがって

$$(1.13) \quad P(X=k)={}_n C_k p^k q^{n-k} \quad (k=0, 1, \dots, n)$$

である。なお、 ${}_n C_k$ は二項係数 binomial coefficient ともよばれ、 $\binom{n}{k}$ と書く場合もある。

(1.12) と二項定理の展開式から

$${}_n C_0 q^n + {}_n C_1 p q^{n-1} + \dots + {}_n C_k p^k q^{n-k} + \dots + {}_n C_n p^n = (p+q)^n = 1$$

が成立する。したがって(1.13)は、 X の確率分布となる条件を満たしている。また(1.13)は二項定理の展開式の一般項であることから、この確率分布を二項分布 binomial distribution とよび、 $Bi(n, p)$ と表す。

さてここで、二項分布の期待値と分散を求めてみよう。定義にしたがって導けるが、ここではより省力的な方法を用いよう。

いまベルヌーイ試行において、 Z_i を、第 i 回目に S が起きたときに 1、起きないときに 0 という値をとる確率変数とする。このとき X は、

$$X=\sum_{i=1}^n Z_i$$

と表現できる。

$$E(Z_i)=1 \cdot p + 0 \cdot q = p$$

に注意すれば、(1.9)より

$$(1.14) \quad E(X) = \sum_{i=1}^n E(Z_i) = np$$

が導ける。つぎに分散を求めよう。まず Z_i の分散は

$$V(Z_i) = (1-p)^2 \cdot p + (0-p)^2 \cdot q = pq$$

となる。 $Z_i (i=1, 2, \dots, n)$ は、独立な確率変数列であるから、(1.11)より

$$(1.15) \quad V(X) = \sum_{i=1}^n V(Z_i) = npq$$

が導ける。

ここで二項分布の形を見てみよう。図1.1(a)は $p=0.5$ に固定して n を変化させたものである。 n が大きくなるとともになめらかな曲線に近づいていく。逆に図1.1(b)は、 $n=10$ に固定して p を変化させたものである。期待値 np の近辺で分布は最大値をとる。また $p=q=0.5$ のとき分布は対称であるが、それ以外の場合は非対称になる。

幾何分布と負の二項分布

ベルヌーイ試行において、初めて事象 S が起きるまでに要した回数を X とする。 X の確率関数を求めよう。

$X=r$ は、ベルヌーイ試行 $Z_i (i=1, 2, \dots)$ において $Z_1=Z_2=\dots=Z_{r-1}=0, Z_r=1$ と同値であるから

$$(1.16) \quad P(X=r) = pq^{r-1} \quad (r=1, 2, \dots)$$

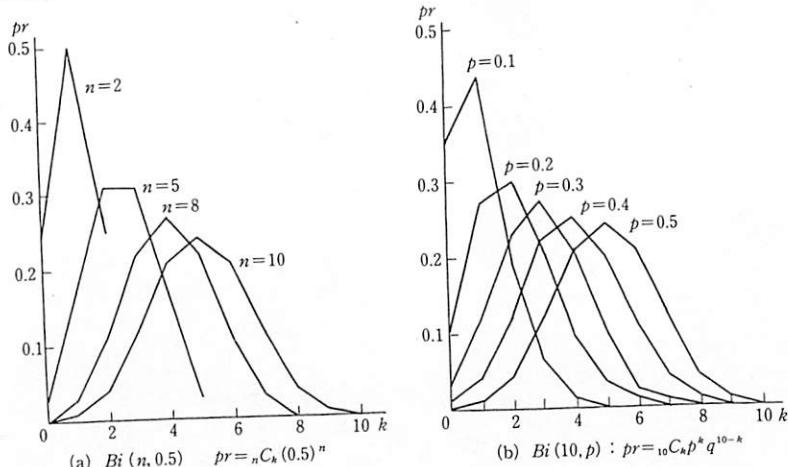


図1.1' 二項分布 $Bi(n, p)$ の形

二項分布はさまざまな分布の基礎となる。 $p=0.5$ のとき対称となり、 n が大きくなるとともに横に広がった滑らかな曲線に近づく。また np の近辺で最大値をとる。

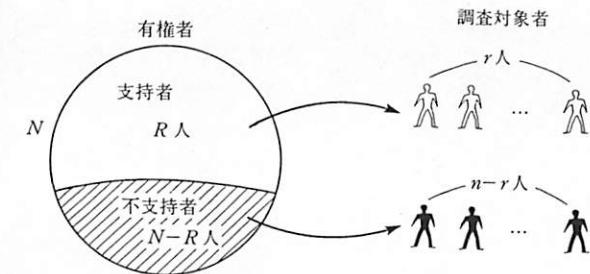


図1.2 超幾何分布

集団を構成する個々について、「支持か不支持か」、「男か女か」、「就業中か失業中か」などある特性を持つか否かに注目したとき、調査した部分の中で、特性を持つものの比率を表す分布である。全体の数が大きくなるにつれて二項分布に近づく。

となる。この確率分布はその式の形から幾何分布 geometric distribution とよばれている。

これを一般化して、ベルヌーイ試行において、事象 S が M 回起きるまでに要した回数を X とする。 X の確率関数を求めよう($M=1$ のときは、幾何分布に一致する)。

$X=n$ は、ベルヌーイ試行 $Z_i (i=1, 2, \dots)$ において、 Z_1, Z_2, \dots, Z_{n-1} のうち $(M-1)$ 個が 1, $(n-M)$ 個が 0, そして $Z_n=1$ と同値である。したがって

$$(1.17) \quad P(X=n) = {}_{n-1} C_{M-1} p^{M-1} q^{n-M} p = {}_{n-1} C_{M-1} p^M q^{n-M}$$

となる。この確率分布は負の二項分布 negative binomial distribution とよばれている。

超幾何分布 ある都市に、 N 人の有権者がいるとしよう。そのうちの R 人 ($0 \leq R \leq N$) が、A という政党を支持しているとする。いま N 人から、でたらめに次々に n 人のひと ($n < N$) を選び、A 党を支持するか、しないかを調査したとする。ただし一度調査したひとは、それ以降の調査では除外する(非復元抽出 sampling without replacement という)。

n 人のうち、A 党を支持した人数を X として、 X の確率分布を考えよう。 N 人から n 人のひとを選ぶ選び方は、 ${}_N C_n$ である。その一つ一つの組合せは等確率 $1/{}_N C_n$ で起きる。一方、 n 人の中に r 人 ($0 \leq r \leq \min(n, R)$) の支持者が含まれる選び方は、図1.2のように R 人の支持者から r 人選ぶ方法 ${}_R C_r$ に、

$N-R$ 人の不支持者から $n-r$ 人を選ぶ方法 ${}_N C_{n-r}$ をかけたものに等しい。
したがって

$$(1.18) \quad P(X=r) = {}_R C_r {}_{N-R} C_{n-r} / {}_N C_n \quad (r=0, 1, \dots, \min(n, R))$$

※世論調査に役立つ確率論※

政党支持率・脳死判定・自衛隊の国連平和維持活動への参加問題・小中学校の週休2日制など国民全体にかかわる問題についての民意を把握するうえで、世論調査は欠かせない。効率的かつ正確な調査を行なうため考案されたのが、確率論にもとづく無作為抽出法(ランダム・サンプリング)である。

調査対象全体から、各個人を等しい確率で無作為(ランダム)に選び出す。通常政治に関する調査では、有権者数約9千万人の中から3千人ほどを抽出する。一見少なすぎるように思えるが、眞の支持率を p として、二項分布で近似すれば標本支持率の分散は $p(1-p)/3000$ である。分散最大の $p=0.5$ のときでも、たかだか 0.8×10^{-4} くらいである。実際には、さらに精度を上げるときでも、たかだか 0.8×10^{-4} くらいである。実際に、さらに精度を上げるために職業・年齢・性別・地域などにより層に分けた、層化無作為多段抽出法が用いられている。

ただし、以上はあくまで「机上」の話であり、現実には回答拒否・集計の手違いなどの調査過程における「非標本誤差」も考慮しなくてはならない。

調査方法は表のように面接調査・電話調査に分かれる。各々一長一短があるが、面接調査は正確性にすぐれ、訪問時のトラブルによる回答拒否などに欠点がある。一方、電話調査は速報性・経費の節約にすぐれるが、質問の正確度に欠ける。なお、男女の回答比率を比較すると通常女性の比率が高い。概して男性の方が社会問題に対して関心が薄いのか? それとも他の理由か? 考えられたい。

面接調査と電話調査の違い

	対象者を選ぶ名簿	選び方	対象者	有効回答率
面接調査	選挙人名簿、住民基本台帳	無作為	特定	あり
電話調査	① 選挙人名簿、住民基本台帳	無作為	特定	あり
	② 電話帳	無作為	世帯内で特定	あり
	③ 電話帳	割り当て法	特定せず	なし ^{*)}

^{*)} ランダム・サンプリングではないので、有効回答率の考え方自体がない。

(表は朝日新聞1991年7月より)

である。この確率分布を超幾何分布 hypergeometric distribution という。ところで、もし毎回の調査対象に以前に調査した人も含めるとすれば X は二項分布 $Bi(n, p)$ ($p=R/N$) に従う。これを復元抽出 sampling with replacement という。 N が十分大きいときには、2つの調査方法の違いは無視できる。つまり N が増加するとともに、超幾何分布は二項分布に近づいていく。

ポアソン分布 日常生活では、ある時間帯のなかで何回起きるかが重要であり、さらに午後1時、2時というような決まった時点に起きるとは限らない現象が数多くある。このような現象を、確率的に表現することを考えよう。

ここでは例として電話番号案内に通話する場合を取り上げよう。事象 S = 「通話がある」とおき、いま単位時間内に、図1.3の黒丸(●)の点で S が起きたとする。単位時間をいくつかの区間に分け、各区間で S が起きたか起きたいかに注目しよう。しかも $1/2, 1/3, \dots, 1/n, \dots$ と、だんだん細かい幅に分割していく。丸印(○)の区間は、 S が少なくとも1回起きた区間である。分割の幅を細かくしていけば、最終的には、2回以上 S が起きた区間はなくなる。そこで n が十分大きいとき、各区間において、 S が2回以上起きた確率は0と仮定する。

さらに n 個の区間に分割したとき、各区間で S が起きる確率を p_n 、起きない確率を $q_n (=1-p_n)$ として、ベルヌーイ試行と見なそう。このとき S が起きる回数は、二項分布 $Bi(n, p_n)$ に従う。ところで単位時間内に S が起きる回数の期待値を λ とおけば、二項分布の期待値は np_n であるから

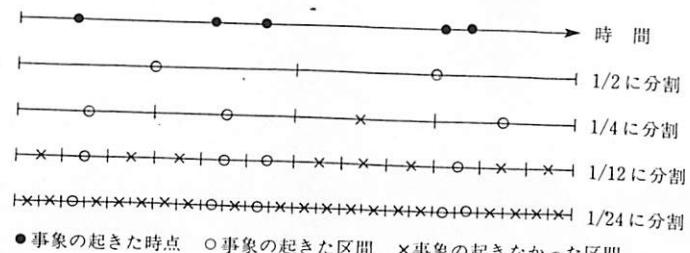


図1.3 ベルヌーイ試行からポアソン分布への移行

二項分布において、時間をだんだん小さい区間に分割し、各区間における生起確率もそれとともに小さくしていった極限がポアソン分布である。これをポアソンの小数の法則ともいう。

$$np_n = \lambda$$

が成立しなくてはならない。したがって、 S が k 回起きる確率は

$$\begin{aligned} & {}_n C_k (p_n)^k (1-p_n)^{n-k} \\ & = {}_n C_k (\lambda/n)^k (1-\lambda/n)^{n-k} \\ & = (\lambda^k/k!) (1-1/n)(1-2/n) \cdots (1-(k-1)/n) (1-\lambda/n)^{-k} (1-\lambda/n)^n \end{aligned}$$

である。ここで $n \rightarrow \infty$ とすれば、右辺の第一項は、 n に無関係であり、最後の項は $e^{-\lambda}$ に収束し、他のすべての項は 1 へ収束する。したがって、極限値として

$$e^{-\lambda} \lambda^k / k!$$

を得る。確率変数 X の確率分布が、

$$(1.19) \quad P(X=k) = e^{-\lambda} \lambda^k / k! \quad (k=0, 1, 2, \dots)$$

となるとき、この分布をパラメータ λ のポアソン分布 Poisson distribution といい、 $P_o(\lambda)$ とあらわす。たしかに

$$e^{-\lambda} \sum_{k=0}^{\infty} \lambda^k / k! = e^{-\lambda} e^{\lambda} = 1$$

であるから、確率分布の条件をみたしている。図 1.4 はパラメータ λ を変化させたときのポアソン分布の形を示している。 λ が大きくなるとともに分布の中心が右側に移動して行き、またそれが広がっていく。

導き方からわかるように、ポアソン分布は、1 回 1 回の試行では起きる確率

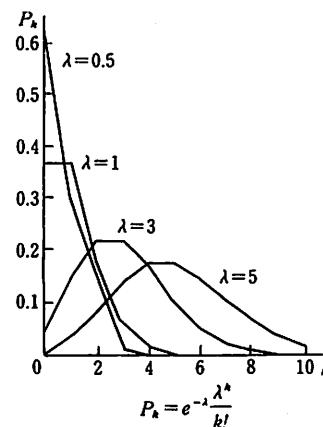


図 1.4[†] ポアソン分布の形状
 λ が大きくなるとともに、分布の中
心が右に移動していく。二項分布と見
なせば $np = \lambda$ であるから、図 1.1 から
想像がつく。

[†]) parameter の訳。「母数」とも訳すが、本章ではより広いこの訳とした。

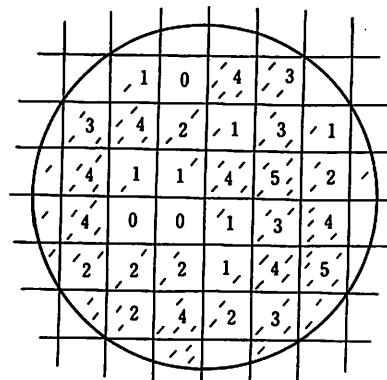


図 1.5 ペトリ板上のバクテリア
生物・植物の生態・繁茂の状況を示す分布
としてもポアソン分布は応用されている。

表 1.1[†] バクテリアの数

k	0	1	2	3	4	5	6 以上
観測度数	5	19	26	26	21	13	8
ポアソン理論値	6.1	18.0	26.7	26.4	19.6	11.7	9.5

(出典：フェラー)

は小さいが、試行回数が多いので全体としてはかなり起こりうる現象に、よくあてはまる分布である。

例 1.1 <バクテリアのカウント数> 図 1.5 はバクテリア集落についているペトリ板を描いたものである。顕微鏡で拡大すると、バクテリア群は黒点に見える。板を小正方形に分割し、ちょうど k 個 ($k=0, 1, 2, \dots$) の黒点をもつ正方形の個数を、度数分布にまとめたのが表 1.1 である。パラメータ λ は未知である。そこで λ が期待値に等しいことに注意して、度数分布の平均値 $\lambda = 2.962$ を用いて計算したのがポアソン理論値である。非常にあてはまりの良いことが見てとれる。

1.1.3 連続分布

この項では統計学への応用上重要な連続分布として、正規分布、指数分布、ガンマ分布、ベータ分布を取り上げよう。

正規分布 さまざまな実験、観測におけるデータや測定誤差のヒストグラムは、観測数を増やすと、頂上が一つ、左右対称、中心から離れるにつれ確

✿確率のきらいな大物理学者もいた✿

近代確率論の創始者 S. ラプラスは次の警句で知られる。「人生の最も重要な問題は、その大部分は、実に確率の問題にすぎない」。物理学者たちは確率をどう考えただろうか。

「誰かが息を吸うとき、その空気の中にジュリアス・シーザーの最後の息の、ある分子がはいっていることは十分にあり得る」

「この一致は、もしそれが偶然性によるならば、その確率は $(1/2)^{60}$ よりもなお小さい。……よってこの一致は、何かの原因によってひき起こされたものにちがいない」

「(量子論は)実在するものの不完全な表現になっている。ただし、表現が不完全であるということは、その規則性の本性が統計的なものであることに必然的に対応している。」

「量子論というものは、将来の物理学に対して有用な基礎を与えるとはとうてい思えない。……それは、量子論的な記述を個々の物理的系または過程の完全な記述と見なそうとすると、矛盾にまきこまれることになるということがある」

最初のものはジーンズ^{*)}によるもの。次のものは、キルヒホップ^{**)}によるもので、彼は太陽スペクトル中の 60 本の線スペクトルが、鉄のそれに一致するところから、太陽中の鉄の存在を推論した。最後の 2 つは、AIN シュタインの、量子力学の確率的(統計的)解釈に対する不満を述べたもの。AIN シュタインと量子力学の関係は從来からよく話題にされる。

^{*)} J. Jeans(1877-1946) 有名な黒体輻射の「レーリー・ジーンズの公式」を導いて前期量子論の先駆をなした。

^{**) G. R. Kirchhoff(1824-1887) ドイツの物理学者。電流に関するいわゆる「キルヒホフの法則」などで有名。黒体輻射の理論、弾性論、音響学、熱学などで多くの著名な業績をあげた。}

率が急速に減少する分布に近づく場合が多い。この 3 つの特徴を兼ね備えた連続分布が、正規分布である。正規分布は、天体観測などにおいて生じる測定誤差の解析から、ガウス(Gauss, C. F.)が考案したという歴史にちなみ、ガウス分布 Gaussian distribution ともよばれている。のちに述べる中心極限定理でも、正規分布は大きな役割をはたす。

いま確率変数 x の確率密度関数が

$$(1.20) \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

で与えられる分布を、正規分布 normal distribution という。ここで π は円周率であり、 e は自然対数の底(約 2.72)である。また、 μ と σ は正規分布の位置(中心)と形を規定するパラメータで、 $-\infty < \mu < +\infty$, $0 < \sigma$ である。簡単のため、この分布を $N(\mu, \sigma^2)$ と記す。N は、normal distribution の頭文字である。また、 μ と σ^2 は正規分布の期待値と分散になっている。とくに $N(0, 1)$ を標準正規分布 standard normal distribution という。

ここで $f(x)$ が密度関数であることを確認しておこう。変数変換 $z = (x - \mu)/\sigma$ をおこなうと

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = 1$$

となる。ここで後の等式はます

$$I = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx$$

とおけば

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dx dy$$

となる。ここで変換 $x = r \cos \theta$, $y = r \sin \theta$ によって極座標 (r, θ) へ移って

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_0^{\infty} e^{-\frac{r^2}{2}} r dr = \left[-e^{-\frac{r^2}{2}} \right]_0^{\infty} = 1$$

より導ける。したがって、たしかに密度関数になっている。

ここで正規分布の実際の形を見よう。(1.20)からわかるように、正規分布の

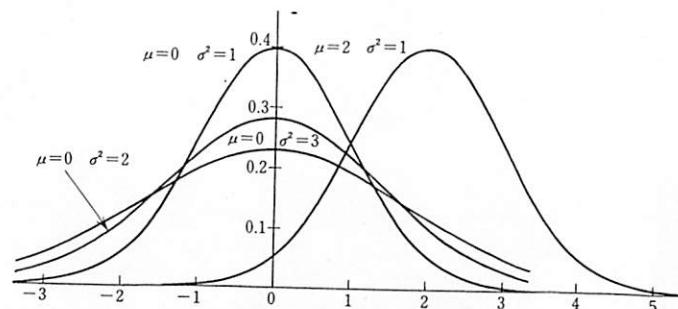


図 1.6 正規分布の密度関数

正規分布の特徴は(1)ひと山型(2)左右対称(3)裾が短い、ということである。 μ が対称の中心であり、 σ が山のなだらかさを決める。

密度関数は、 μ, σ によって決定される。図1.6には、 μ, σ をさまざまな値に変化させたときの、密度関数の変化を示してある。教会のベルのような形をしているが、次のような特徴がある。まず期待値 μ を中心にして左右対称であり、期待値 μ の値が変化すると、形は同じままで平行移動する。分散 σ^2 が大きくなるにつれて、なだらかな山の形になり、 $x=\mu$ から遠くはなれた値をとる確率が大きくなる。

具体的に X がある区間に入る確率を求めるためには、密度関数を積分しなくてはならない。しかしいちいち計算するのは面倒である。そこで標準正規分布については、あらかじめさまざまな区間の積分の値を計算した正規分布表が用意されている。標準正規分布の密度関数と累積分布関数をそれぞれ

$$(1.21) \quad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$(1.22) \quad \Phi(z) = P(X \leq z) = \int_{-\infty}^z \phi(x) dx$$

と表す^{*)}。密度関数が原点に関して対称であることに注意すれば

$$\Phi(-z) = 1 - \Phi(z)$$

という関係が成立する。したがって、正の z に対する $\Phi(z)$ の値がわかれば十分である。通常、 $z \geq 0$ に対する $\Phi(z)$ あるいは、上側の確率

$$Q(z) = 1 - \Phi(z)$$

の表が用意されている。図1.7の斜線部の面積が $\Phi(z)$ を、白い部分の面積が $Q(z)$ を表す。なお、 $0 < \alpha < 1$ に対し、 $Q(z) = \alpha$ となる点 z を z_α と表すことがある。

例1.2 〈正規確率の計算〉 X が $N(0, 1)$ に従うとき、 $X \leq 1.18$ となる確

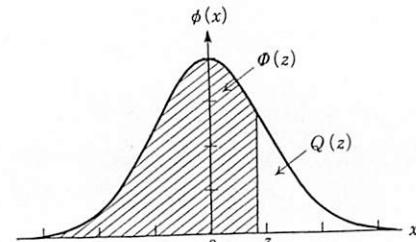


図1.7[†] 標準正規分布の確率
標準正規分布では、斜線部が z 以下の値をとる確率、白い部分が z 以上の値をとる確率である。通常どちらか一方の数表が掲載されている。

^{*)} ϕ と Φ は、それぞれギリシャ文字ファイの小文字と大文字である。

率は

$P(X \leq 1.18) = \Phi(1.18) = 1 - Q(1.18) = 0.88100$
となる^{*)}。また、区間 $[-1.11, 1.17]$ の確率は

$$\begin{aligned} P(-1.11 \leq X \leq 1.17) &= \Phi(1.17) - \Phi(-1.11) \\ &= 1 - Q(1.17) - Q(1.11) \\ &= 0.74550 \end{aligned}$$

である。

次に X が $N(\mu, \sigma^2)$ に従うときを考えよう。まず Y を

$$(1.23) \quad Y = (X - \mu) / \sigma$$

と定義する。これから、期待値と分散は

$$\begin{aligned} E(Y) &= E(X - \mu) / \sigma = (\mu - \mu) / \sigma = 0 \\ V(Y) &= V(X) / \sigma^2 = 1 \end{aligned}$$

となり、実は Y も正規分布に従っている（分布関数あるいは2節のモーメント母関数を用いて証明できるが、原点を平行移動させて目盛りの尺度を変えていいだけであるから、密度関数の形を考えれば、容易に想像がつく）。 X から Y への変換は、一般の正規分布 $N(\mu, \sigma^2)$ から標準正規分布 $N(0, 1)$ への変換であるから、標準化とよばれている。たとえば、 X が区間 $[a, b]$ に値をとる確率は

$$P(a \leq X \leq b) = P((a - \mu) / \sigma \leq Y \leq (b - \mu) / \sigma)$$

と変形すれば計算できる。

例1.3 〈テストのランキング〉 通常、学力テストで用いられる偏差値は、素点を標本平均=50、標本標準偏差=10となるように変換したものである。テストの得点の分布が正規分布に従うと見なして、素点を正規分布 $N(50, 10^2)$ に変換したとするならば、ある学生のテストの偏差値が63のとき

$$P(X \geq 63) = P((X - 50) / 10 \geq 1.3) = 0.0968$$

となり、この学生は、この点数の起り易さの確率から、上位約10%に位置していることがわかる。ただし正規分布と見なせない場合もあるので、偏差値のみを見て、一喜一憂してはいけない。

ところで標準化の手続きからわかるように、 σ で割れば、すべての正規分布

[†]) $\Phi(z)$ を表示する表が多いが、本書の巻末の正規分布表は $Q(z)$ を表示する。

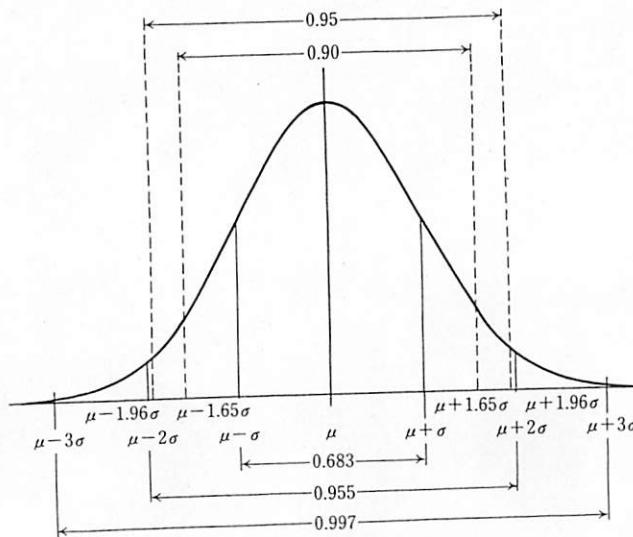


図 1.8 正規分布の基本的な区間確率

正規分布では $\pm 3\sigma$ を超えることは、「千に三つ」も起きない。「まゆつば」だと思った方が無難である。また「 $\pm 1.65\sigma$ 以内は 90%， $\pm 1.96\sigma$ 以内は 95%」は頻繁に使うので、暗記しておいた方が便利である。

は、分散 1 の正規分布に変換できる。したがって σ は、尺度の単位と解釈できる。そこで期待値 μ を中心として、 σ の倍数で区間を分割し、 $\mu \pm \sigma$, $\mu \pm 2\sigma$, $\mu \pm 3\sigma$ 、それぞれの範囲に、 X が値をとる確率を求めるとき図 1.8 のように μ から両側に離れるにつれて、確率は急速に減少する。ちなみに品質管理などでは、 $\mu \pm 3\sigma$ の外側に値が生じたとき、千に三つも起きないことから、工程に異常が生じていると判断される場合が多い。また $\mu \pm 1.65\sigma$, $\mu \pm 1.96\sigma$ の範囲内に値をとる確率は、それぞれ 0.90, 0.95 と、切りのよい数字になる。

指数分布 指数分布は、ポアソン分布に従って起きる事象の生起間隔を表現する分布である。指数分布は、さまざまなサービスの待ち時間を表現する分布として用いられる。

例 1.4 <サービスの待ち時間> いま駅の切符売り場で、前売り乗車券を買ふために、列をなして待っている場面を考えよう。単位時間を n 等分して、 k 番目の区間で、誰かが購入を終了したとき、 $X_k=1$ 、そうでないとき $X_k=0$ とする。切符発売開始時点から計測を始め、列の最初にいた人が購入を終了し

た時点を T とする。 X_1, X_2, \dots は、パラメータ $p=\lambda/n$ の、ベルヌーイ試行に従うとして、 T の確率分布を求めよう。 $N+1$ 番目の区間で購入を終えたとすれば、 n が十分大きいとき、 $T \approx (1/n)N$ と、近似できる。任意の $t (>0)$ に対して、 k を、 $k \leq nt < k+1$ を満たす整数とする。このとき、 $N+1$ が幾何分布に従うことから

$$\begin{aligned} P(0 \leq T \leq t) &\equiv P(0 \leq N \leq nt) \\ &= P(N=0) + P(N=1) + \cdots + P(N=k) \\ &= p + pq + \cdots + pq^k \\ &= 1 - q^{k+1} \\ &\approx 1 - (1 - \lambda/n)^{nt} \end{aligned}$$

ここで $n \rightarrow \infty$ とすれば、右辺は

$$1 - e^{-\lambda t} = \int_0^t f(x) dx$$

に収束する。ただし $f(x)$ は

$$(1.24) \quad \begin{aligned} f(x) &= \lambda e^{-\lambda x} && (x \geq 0 \text{ のとき}) \\ &= 0 && (x < 0 \text{ のとき}) \end{aligned}$$

とする。

そこで(1.24)の $f(x)$ を密度関数にもつ確率分布を、指数分布 exponential distribution といい、 $Ex(\lambda)$ と表す。なお、最初の人が購入した時点を始点として新たに計測を開始したとすれば、二番目の人が購入を終了するまでに要する時間も同じ分布に従う。以下同様である。

ガンマ分布 密度関数

$$(1.25) \quad f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad (x \geq 0)$$

を持つ分布をガンマ分布 Gamma distribution といい、 $Ga(\alpha, \lambda)$ と表す^{*}。ここで、関数 $\Gamma(\alpha)$ はガンマ関数 Gamma function で、 $\alpha > 0$ に対し、

$$(1.26) \quad \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

^{*}) x の指数部分を x/λ と書く方式を $Ga(\alpha, \lambda)$ と書く記法もある。この記法では、尺度パラメータが λ そのものとなる。

$$M^{(1)}(0)=E(X)=\mu, \quad M^{(2)}(0)=E(X^2)=\sigma^2+\mu^2$$

だから、これから $V(X)=\sigma^2$ が導ける。

独立な確率変数の和の分布 X, Y を互いに独立な確率変数とし、 $M_X(t), M_Y(t)$ をそれぞれのモーメント母関数とする。また $M_{X+Y}(t)$ を $X+Y$ のモーメント母関数とする。このとき e^{tX} と e^{tY} も互いに独立であるから

$$\begin{aligned} M_{X+Y}(t) &= E(e^{t(X+Y)}) = E(e^{tX})E(e^{tY}) \\ (1.35) \quad &= M_X(t)M_Y(t) \end{aligned}$$

となる。したがって、独立な確率変数の和のモーメント母関数は、それぞれの確率変数のモーメント母関数の積になる。

いま X, Y が互いに独立な確率変数で、それぞれ正規分布 $N(\mu_X, \sigma^2_X)$, $N(\mu_Y, \sigma^2_Y)$ に従うとする。このとき (1.31), (1.35) より

$$M_{X+Y}(t) = \exp \{(\mu_X + \mu_Y)t + (\sigma^2_X + \sigma^2_Y)t^2/2\}$$

となる。モーメント母関数と確率分布は 1 対 1 に対応することが知られているので、 $X+Y$ の確率分布もやはり正規分布 $N(\mu_X + \mu_Y, \sigma^2_X + \sigma^2_Y)$ になる。この性質は再生性とよばれている。このように独立な確率変数の和の分布を導くのにも、モーメント母関数は有効な道具となる。

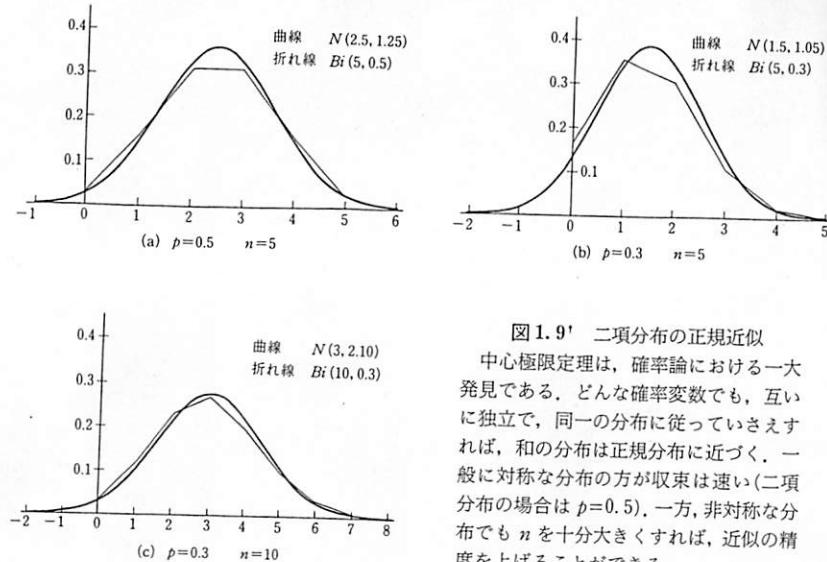
1.3 中心極限定理

二項分布 $Bi(n, p)$ の形状を示した図 1.1(a) を再度ながめてみよう。 n を大きくすると、だんだんなめらかで連続な分布に近づいていく様子がわかる。実際には n の増加とともに分布関数は正規分布に近づいていく。二項分布は、 X_i ($i=1, 2, \dots, n$) をベルヌーイ試行列とすれば、和

$$(1.36) \quad S_n = X_1 + X_2 + \dots + X_n$$

の分布と同じであった。しかし二項分布に限らず、 X_i ($i=1, 2, \dots, n$) が、互いに独立で同一の確率分布に従う場合、 S_n の分布は、常に正規分布に近づいていくことが知られている。これを中心極限定理 central limit theorem とよび、さまざまな分布の近似として正規分布を用いることの一つの正当化である。きちんと述べると、次のようになる^{*)}。

*) 簡単な証明は本シリーズ I『統計学入門』第 8 章で与えられている。



(a) $p=0.5, n=5$
曲線 $N(2.5, 1.25)$
折れ線 $Bi(5, 0.5)$

(b) $p=0.3, n=5$
曲線 $N(1.5, 1.05)$
折れ線 $Bi(5, 0.3)$

(c) $p=0.3, n=10$
曲線 $N(3.2, 1.10)$
折れ線 $Bi(10, 0.3)$

図 1.9† 二項分布の正規近似
中心極限定理は、確率論における一大発見である。どんな確率変数でも、互いに独立で、同一の分布に従っていさえすれば、和の分布は正規分布に近づく。一般に対称な分布の方が収束は速い(二項分布の場合は $p=0.5$)。一方、非対称な分布でも n を十分大きくすれば、近似の精度を上げることができる。

X_i ($i=1, 2, \dots, n$) が、互いに独立かつ同一の確率分布に従う確率変数列で、 $E(X_i)=\mu, V(X_i)=\sigma^2 (<\infty)$ とする。このとき

$$(1.37) \quad P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) \rightarrow \Phi(x) \quad (n \rightarrow \infty)$$

が成り立つ。

例 1.6 〈二項分布の正規近似〉 二項分布 $Bi(n, p)$ で $p=0.3, 0.5$ の場合を見よう。左右対称な分布である $p=0.5$ の方が、非対称な $p=0.3$ より n が同じならば、近似の精度がよいことがわかる(図 1.9(a), (b))。ただ $p=0.3$ の方も、 n をさらに大きくすれば精度は上がる(図 1.9(c))。

n が大きいとき、真の分布を用いて確率を計算するのは面倒な場合が多い。そのときには、正規分布で近似して確率を求めた方がはるかに効率的である。一例として、さいころを 120 回投げて 1 の目が 25 回以上出る確率を求めてみよう。真の分布は $Bi(120, 1/6)$ であるが、正規分布で近似すれば

$$P\left(\frac{S_{120} - 20}{\sqrt{120 \cdot \frac{1}{6} \cdot \frac{5}{6}}} \geq \frac{25 - 20}{\sqrt{120 \cdot \frac{1}{6} \cdot \frac{5}{6}}}\right) \approx Q(1.22) = 0.11$$

となる。期待値は $120 \cdot (1/6) = 20$ だから、5だけ上まわる確率はここまで低下することがわかる。

◆◆◆◆ 第1章 練習問題 ◆◆◆◆

1.1 <二項分布> 二項分布 $Bi(n, p)$ の期待値と分散を、定義に従って導け。

1.2 <幾何分布> 幾何分布に従う確率変数 X の期待値を求めよ。

1.3 <負の二項分布> 負の二項分布に従う確率変数の期待値を求めよ。

1.4 <超幾何分布> 超幾何分布の確率関数を

$$p_r = {}_R C_r N - {}_R C_{n-r} / {}_N C_n$$

とおく。

i) $K = \min(n, R)$ として

$$\sum_{r=0}^K p_r = 1$$

を示せ。

ii) $p = R/N$ を固定したまま $N \rightarrow \infty$ とすると、二項分布 $Bi(n, p)$ に近づくことを確かめよ。

1.5 <ポアソン分布> ポアソン分布 $Po(\lambda)$ の期待値と分散を、定義に従って導け。

1.6 <続、あてはめ> i) 表1は、1875年から1894年までの20年間、毎年10個の隊で(延べ200の隊)、馬に蹴られて死んだプロシャの兵士の数を集計したものである。

ポアソン分布に従うとみなし、あてはまり具合を調べよ。ただしパラメータ λ には、実測値の平均値を代入せよ。

表1'	1年間の死者	0	1	2	3	4	計
	隊の数	109	65	22	3	1	200

注) ポーランド系のドイツの経済学者ボルトキーヴィッヂ(Bortkiewicz, L. v.)がまとめた。

ii) 表2は、ある都市の1年間の毎日の交通事故による死亡者数である。i)と同じ計算を行え。もしポアソン分布のあてはまりが良くない場合はその原因を考えよ。

表2	死者	0	1	2	3	4	5	6	7	計
	日数	74	98	90	61	21	13	7	1	365

1.7 <一様分布> 一般に

$$f(x) = \begin{cases} 1/(b-a) & (a < x < b) \\ 0 & (x \leq a, x \geq b) \end{cases}$$

1.3 中心極限定理

(ただし $a, b (a < b)$ は定数) を密度関数にもつ分布を、一様分布という。一様分布の期待値と分散を求めよ。

1.8 <ポアソン確率変数の和> X_1, X_2 が互いに独立な確率変数で、各々パラメータ λ_1, λ_2 のポアソン分布に従うとする。 $X_1 + X_2$ のモーメント母関数を求めよ。これから $X_1 + X_2$ は、パラメータ $\lambda_1 + \lambda_2$ のポアソン分布に従うことを示せ。

1.9 <指数分布とガンマ分布> $X_i (i=1, 2, \dots, n)$ は互いに独立で、同一の指数分布 $Ex(\lambda)$ に従うとする。

$$S_n = X_1 + X_2 + \dots + X_n$$

と定義する。

i) 和 S_n のモーメント母関数を求めよ。

ii) ガンマ分布 $Ga(\alpha, \lambda)$ のモーメント母関数を求めよ。

iii) S_n の分布は、ガンマ分布 $Ga(n, \lambda)$ に等しいことを示せ。

1.10 <ポアソン分布と中心極限定理> $X_i (i=1, 2, \dots, n)$ は互いに独立で、パラメータ $\lambda=1$ の同一のポアソン分布に従うとする。中心極限定理を適用して

$$e^{-n} \sum_{k=0}^n n^k / k! \rightarrow 1/2 \quad (n \rightarrow \infty)$$

を示せ。

第2章 線形モデルと最小二乗法

統計的推測はモデルを通して行う場合とモデルを用いない場合があり、それぞれパラメトリック、ノンパラメトリックとよばれる。そして、パラメトリックの代表選手が期待値に線形な構造を、誤差の確率分布に正規分布を仮定する正規線形モデルである。分散分析や回帰分析の理論はすべてこの線形モデルに基づいているので、本章の基礎理論がわかると多くの具体的な問題を解析することができる。

2.1 線形モデル

2.1.1 線形モデルの導入

統計的推論は誤差に埋もれたデータから、系統的な意味のある情報を引き出す科学的プロセスということができる。そのためにパラメトリックあるいはノンパラメトリックな様々な手法が用いられる。その中で、観測値 y に

未知母数の線形式+誤差

という構造を仮定する線形モデル linear model が分散分析や回帰分析の基本モデルとして幅広く用いられている。

いま、そのもっとも簡単な例として、ある物体の重さ μ を計測するモデル
(2.1)
$$y = \mu + \varepsilon$$

を考えよう。ここで ε は、たとえば操作あるいは観測などに伴う偶然誤差である。 μ が一定でも観測のたびに観測値 y が異なるのはこれら計測誤差が避けられないためである。ところで、 μ をただ1回計測したのでは μ と ε を分

離する術がない。そこで計測を独立に n 回繰返すことにすると、繰返し測定のモデル

$$(2.2) \quad y_i = \mu + \varepsilon_i, \quad i=1, \dots, n$$

が得られる。このようにしても、相変わらず n 個の測定値に対し未知数は $n+1$ 個あるので μ を正確に(単独で)知ることはできない。そこで個々の誤差を知ることはあきらめて、確率的なふるまいを推定することにする。

そのときよく用いられるのは ε_i がたがいに独立に期待値(平均)0、分散 σ^2 の正規分布 $N(0, \sigma^2)$ に従うという仮定である。正規分布は期待値と分散で完全に規定され、たとえば期待値のまわり $\pm\sigma$ 内の確率は 0.682, $\pm 2\sigma$ 内は 0.955、そして $\pm 3\sigma$ 内は 0.997 であることなどがよく知られている。とくに $\pm 3\sigma$ 外のデータの生じる割合が 3/1000 ということは正規分布が比較的おとなしい分布であることを意味しているが、それは $N(\mu, \sigma^2)$ の密度関数

$$(2.3) \quad f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}$$

が、 y が μ から離れるに従い、急速に減衰するからである。もし、正規分布が仮定できれば、 σ^2 を推定することにより、誤差のふるまいが推定できることになる。このように、誤差に正規性を仮定する線形モデルはとくに正規線形モデル normal linear model とよばれる。

2.1.2 真の値 μ と誤差 ε_i の分離

さしあたり、誤差 ε_i は正規分布 $N(0, \sigma^2)$ に、したがって y_i は $N(\mu, \sigma^2)$ に従うとしておこう。(この仮定の妥当性については 2.1.5 で述べる)。このように ε_i を確率的誤差と考えると、どうして μ と ε が分離できるのだろうか。

\times …生データ (▲は×データの平均)
 \circ …10個ずつの平均 (△は総平均)

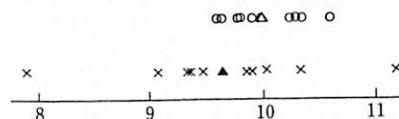


図 2.1 正規乱数のプロットによる数値実験
▲の方よりも△の方が真の平均 10 に近く、
平均をとった後者の方が集中の傾向が見られる。

ここで数値実験をしてみよう。いま、 $N(10, 1^2)$ に従う乱数を独立に 10 個発生させたものを図 2.1 上に \times で示す。同じような実験をさらに 9 回繰返し、各回の平均のみを図 2.1 上に \circ で示す。同じく 10 個のデータでありながら、明らかに平均の方が散らばり具合が小さく、かつ横軸 10 の近くに集まっている。さらに全体(100 個)の平均は 10.027(△で示した)となり、真の平均 10 にきわめて近い値が得られている。

これは数学的に次のように説明することができる。いま、 y_1, \dots, y_n がたがいに独立に正規分布 $N(10, 1^2)$ に従って分布しているとする。このとき平均 \bar{y} はまた正規分布に従い、期待値は

$$\begin{aligned} E(\bar{y}) &= E\{(y_1 + \dots + y_n)/n\} \\ &= \{E(y_1) + \dots + E(y_n)\}/n = 10 \end{aligned}$$

分散は

$$\begin{aligned} V(\bar{y}) &= V\{(y_1 + \dots + y_n)/n\} \\ &= \{V(y_1) + \dots + V(y_n)\}/n^2 = 1/n \end{aligned}$$

となる。すなわち、図 2.2 に示すように中心は同じまま平均の分布の方がばらつきが小さくなるのである。

(2.2) のようなモデルで ε_i がたがいに独立に $N(0, \sigma^2)$ に従うと仮定し、繰返し数 n を大きくとると、 \bar{y} の分布は $N(\mu, \sigma^2/n)$ となって μ の近くに集中していくことがわかった。既に述べたように \bar{y} が μ から、たとえば $\pm 3\sqrt{\sigma^2/n}$ 以上離れることは 3/1000 の確率でしか生じない。そして n を大きくすれば $3\sigma/\sqrt{n}$ はいくらでも小さくできるのである。これで正規分布の場合に測定を繰返し、 \bar{y} を計算することによって μ を精度よく推定できることがわかった。

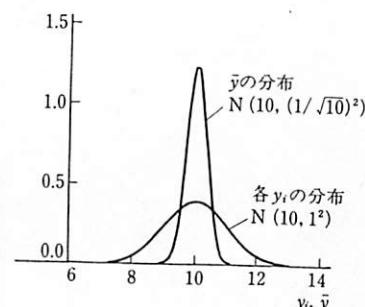


図 2.2 平均 \bar{y} の分布とともに各 y_i の分布との関係
平均 \bar{y} の分布は各 y_i の分布と同じく真の平均 μ を中心しているが、 μ への集中の傾向は繰返し数 n が大きくなるにつれ強くなる。

2.1.3 大数の弱法則

2.1.2と似たことは正規分布の仮定なしにも示すことができる。いま、 $\varepsilon_1, \dots, \varepsilon_n$ がたがいに無相関で期待値0, 分散 σ^2 の分布に従っているとすると、先と同じように

$$E(\bar{y}) = \mu, \quad V(\bar{y}) = \sigma^2/n$$

が成り立つ。このときチエビシェフの不等式によって

$$(2.4) \quad P(|\bar{y} - \mu| \geq \delta) \leq \frac{\sigma^2}{n\delta^2}$$

が成り立つ。ここで δ がどんなに小さな正定数であっても、 n を十分大きくすることにより、(2.4)の右辺はいくらでも小さくすることができる。すなわち、 n を大きくすると \bar{y} が μ の近くに集中してくることが2次モーメント(積率)までの仮定でいえる。この事実は大数の弱法則(簡単に大数の法則)とよばれる。

2.1.4 推定量の良さ

実は n とともに μ の近くに集中してくる推定量は \bar{y} の他にいくらでも考えられる。たとえば y_1, \dots, y_n の中央値 y_{med} やホッジス・レーマン推定量 Hodges-Lehmann estimator

$$y_{HL} = \left(\text{あらゆる } \frac{y_i + y_j}{2} \text{ (ただし } i \leq j \text{) の中央値} \right)$$

などもその性質を満たしている。その中で \bar{y} はどういう意味で良い推定量なのだろうか。

いま、モデル(2.2)で、誤差 ε_i は性質

- i) 不偏性: $E(\varepsilon_i) = 0$
- (2.5) ii) 等分散性: $V(\varepsilon_i) = \sigma^2$
- iii) 無相関性: $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (i \neq j)$

を満たす確率変数であるとする。

ここで y_i の線形結合で表される不偏*の推定量

$$l_1 y_1 + l_2 y_2 + \dots + l_n y_n$$

*) 推定量の不偏性は、本シリーズ I『統計学入門』第9章、第11章参照。

2.1 線形モデル

の中で最小分散のものを考えてみよう。まず、誤差についての仮定(2.5)を用いて計算すると

$$V(\sum l_i y_i) = \sum V(l_i y_i) = (\sum l_i^2) \sigma^2$$

であることがわかる。一方、不偏性の条件

$$E(\sum l_i y_i) = \mu$$

は

$$E(\sum l_i y_i) = \sum E(l_i y_i) = \sum l_i \mu = \mu$$

と表され、これが任意の μ について成り立たねばならないから

$$\sum l_i = 1$$

となるべきである。そこで、量 $\sum l_i^2$ を条件 $\sum l_i = 1$ の下で最小にすると

$$l_1 = l_2 = \dots = l_n = 1/n$$

が解となり

$$\sum l_i y_i = \sum y_i / n = \bar{y}$$

が得られる。すなわち、 \bar{y} は線形な推定量に限って考えれば、誤差に関する仮定(2.5)の下で最小分散の不偏推定量であることがいえる。このような推定量を最良線形不偏推定量 Best Linear Unbiased Estimator あるいは BLUE とよぶ。もっと一般的な線形モデルで、いちいちこのような分散最小化問題を解くことなく BLUE を得ようというのがこれから述べる最小二乗法である。

実は、証明は省くが、誤差についての仮定(2.5)に加えて、その分布が正規分布であると仮定すると(また、そのときに限って)、 \bar{y} は線形に限らずあらゆる関数形の不偏推定量の中で最小分散であることがいえる。一方、誤差 ε の分布がコーシー分布

$$f(\varepsilon) = \frac{1}{\pi} \cdot \frac{1}{1 + \varepsilon^2}, \quad -\infty < \varepsilon < \infty$$

のように期待値、分散が存在しないような分布では、2.1.2節からここまで述べてきたような \bar{y} の最適性は成り立たず、それに替って中央値 \tilde{y} やホッジス・レーマン推定量 y_{HL} の方が良い推定量となる。コーシー分布は正規分布に比べて裾が長く、ときどき非常に大きなあるいは小さな値が生じる。 \bar{y} のようにすべての観測値を対等に扱う推定量はその極端な値に引きずられるのに対し、その数値を直接使わない y_{med} , y_{HL} の方が安定しているのである。

2.1.5 正規分布の仮定

誤差分布として正規分布を仮定するのはそれほど非現実的なことではない。いま、誤差 ε がきわめて小さく、かつたがいに独立な偶然誤差 $\delta_1, \dots, \delta_n$ の寄せ集めで生じたというモデル

$$\varepsilon = \delta_1 + \dots + \delta_n$$

を考える。ここで各 δ_i がまったく偶然に等確率 $1/2$ で δ または $-\delta$ をとるとすると

$$E(\delta_i) = (1/2) \cdot \delta + (1/2) \cdot (-\delta) = 0,$$

$$V(\delta_i) = (1/2) \cdot \delta^2 + (1/2) \cdot (-\delta)^2 = \delta^2$$

となるので

$$E(\varepsilon) = 0, \quad V(\varepsilon) = n\delta^2$$

となる。しかるに、 ε の分散が無限大に発散するというのは現実のモデルとして不合理なので

$$(2.6) \quad n\delta^2 = \sigma^2 \quad (\text{定数})$$

とおく。次に ε の極限分布を調べるためにモーメント母関数

$$E(e^{t\varepsilon}) = E(e^{t\delta_1})E(e^{t\delta_2}) \cdots E(e^{t\delta_n}) = \{(1/2)e^{t\delta} + (1/2)e^{-t\delta}\}^n$$

において指數関数を展開し、(2.6)に注意すると

$$E(e^{t\varepsilon}) = \left\{ 1 + \frac{t^2}{2} \cdot \frac{\sigma^2}{n} + o\left(\frac{1}{n}\right) \right\}^n \xrightarrow{n \rightarrow \infty} e^{\sigma^2 t^2 / 2}$$

が得られる*). ところがこれは正規分布 $N(0, \sigma^2)$ のモーメント母関数に他ならない。すなわち、誤差がたがいに独立で小さな偶然変動がたくさん寄せ集まつて生じたという仮定が合理的な場では、正規分布の仮定は自然なのである。

もともとたがいに独立に期待値 0、分散 σ^2 に従う変数 y_1, \dots, y_n に関して

$$(y_1 + \dots + y_n) / \sqrt{n}$$

の分布は、 n を大きくすると元の分布が何であっても正規分布 $N(0, \sigma^2)$ に近くことは中心極限定理としてよく知られている。先の大数の法則では n 個の和を n で割っているのに対し、ここでは \sqrt{n} で割っていることに注意して欲しい。 n を大きくする目安は元の分布があまり裾が長くなく、かつ歪んでいないければ 10 程度よい。例として元の分布が区間 $[0, 10]$ の一様分布に従うと

*) $o(1/n)$ は $1/n$ よりも高次の無限小を表す数学記号。…で表しているのを厳密に表現する。

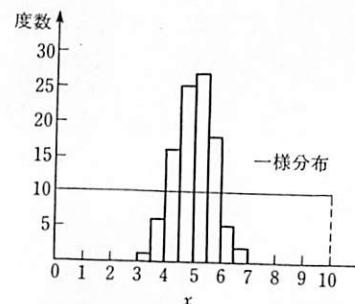


図 2.3 一様乱数 12 個の平均のヒストグラム
一様乱数 12 個 y_1, y_2, \dots, y_{12} の平均 \bar{y} の分布は、中心極限定理から、正規分布に酷似する。この図でもきれいな正規分布があらわれており、正規乱数の手頃な発生法として用いられる。(第 12 章「乱数の知識」参照)

き、12 個ずつの平均 100 個から構成したヒストグラムを図 2.3 に示した。元の一様分布とは似ても似つかない滑らかな分布が得られる。これは一様乱数を元に正規乱数を発生させる方法の一つとして知られている。

2.1.6 行列による表現

線形モデル linear model は分散分析、回帰分析の基本モデルである。たとえば、ある処理(工法、触媒、原料の種類など)の a 通りの水準について、それぞれ r 回の実験を無作為化して行い、各処理効果 α_i を比較するための 1 元配置 one-way layout のモデル*)

$$(2.7) \quad y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i=1, \dots, a; \quad j=1, \dots, r$$

や、温度、電流などの計量要因 x に関する多項式 polynomial モデル

$$(2.8) \quad y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, \quad i=1, \dots, n$$

などは、みな、線形モデルである。

繰り返し測定のモデル(2.2)における μ 、1 元配置のモデル(2.7)における $\mu + \alpha_i$ 、多項式モデル(2.8)における $\beta_0 + \beta_1 x_i + \beta_2 x_i^2$ はそれぞれデータの期待値にあたり、データの系統的な変動を説明する部分である。モデル(2.2)、(2.7)、(2.8)はこれら期待値に仮定(2.5)を満たす誤差がつけ加わったものが観測されることを表している。 μ, α, β が期待値を定める未知母数(パラメータ)で、「線形モデル」というより方は、期待値がこれら未知母数の線形結合で表されるこからきている。したがって、説明変数に x^2 のような非線形項が混じるのは一向に構わない。それはデータを得た後は単なる既知係数だからである。なお、以降線形モデルというより方は誤差についての仮定(2.5)まで含めて用いるこ

*) くわしくは、第 3 章 3.2 節を見よ。

とにする。

ベクトル、行列による表現 線形モデルにおける未知母数の推定・検定あるいは期待値を表現する線形モデルの選択などの理論を総称して線形推測論 linear statistical inference という。その具体的な手法が最小二乗法といふわけである。線形モデルの具体的な取り扱い、および線形推測論の理解には行列を用いるのが便利である。そこで線形モデルを行列で表すことから始めよう。

慣れないうちはやや面倒だが、多数の変数を表現する方法として、むしろベクトル、行列による表現を用いることが便利であることが、慣れるに従いわかってくる。

いま、観測値のベクトルを列ベクトルで

$$\mathbf{y} = (y_1, \dots, y_n)',$$

未知母数のベクトルを一般的に

$$\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)',$$

既知係数行列を \mathbf{X} 、そして誤差ベクトルを

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$$

とする^{*)}と、線形モデルは一般に

$$(2.9) \quad \mathbf{y} = \mathbf{X} \boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

と表すことができる。 n 行 p 列の行列 \mathbf{X} は実験の計画によって定まる係数行列なので、とくに計画行列 design matrix とよばれる。 \mathbf{X} の具体的な形は次の例 2.1～2.3 を参照して欲しい。

例 2.1 〈繰返し測定のモデル〉 (2.2) を線形モデル (2.9) の形式で表すと

$$(2.10) \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

のようになる。すなわち、この場合 \mathbf{X} は要素がすべて 1 の列ベクトルである。

*) 列ベクトルの記法は印刷上スペースを多く必要とするので、まず行ベクトルで書き、それを転置('あらわす)するという便宜法を用いるのがふつうである。'はプライムと読む。

表 2.1 繰返し測定のデータ：ある飲料製品中のエキス含有量(%)

1	2	3	4	5	6	7	8	9
37.55	37.74	37.69	37.71	37.73	37.81	37.53	37.58	37.74(%)

繰返し測定のモデルが典型的にあてはまる例である。

(出典：広津(編) (1993))

なお

$$\mathbf{j}_n = (1, \dots, 1)'$$

すべての要素が 1 である $n \times 1$ 列ベクトルを表すと、たとえば (2.10) は

$$\mathbf{y} = \mathbf{j}_n \boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

のように表される。ただし、次元を表す添字 n は自明な場合には省略する。

表 2.1 のデータは、ある飲料製品中のエキス含有量 % を繰り返し測定した例である。9 個のデータは物理的に違うものだが、みな同じ正規母集団 $N(\mu, \sigma^2)$ からのランダム・サンプルと仮定できるなら繰返し測定のモデルとして表せる。ただし、誤差 $\boldsymbol{\varepsilon}$ はサンプリング誤差と測定誤差の合成からなる。

例 2.2 〈1 元配置モデル〉 (2.7) は

$$(2.11) \quad \mathbf{y} = \begin{bmatrix} 1 & 1 & & & O \\ 1 & 1 & \ddots & & \\ 1 & & 1 & \ddots & \\ \vdots & & & \ddots & \\ 1 & & & & 1 \\ 1 & & & & O \\ 1 & & & & \vdots \\ 1 & & & & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \end{bmatrix} + \boldsymbol{\varepsilon}$$

$$= \begin{bmatrix} j_r & & & & O \\ j_{ar} & \ddots & & & \\ & & \ddots & & j_r \\ O & & & \ddots & j_r \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_a \end{bmatrix} + \boldsymbol{\varepsilon}$$

のようになる。ただし、

$$\mathbf{y} = (y_{11}, y_{12}, \dots, y_{ar})'$$

である。なお、複数添字のデータをベクトルで表すときは一般に添字に関して辞書式に並べるものとする。したがって、線形モデル (2.11) で、 $r(i-1)+j$

表2.2' 1元配置のデータ：脂肪の種類に対応する吸収量

脂 肪	1	2	3	4	計
	64	78	75	55	
	72	91	93	66	
	68	97	78	49	
	77	82	71	64	
	56	85	63	70	
	95	77	76	68	
和	432	510	456	372	1,770
平 均	72	85	76	62	295

総平均: $1770/24=73.75$

(出典: スネデカーーコクラン)

1元配置のモデルは、線形モデルとして最もわかり易い例である。脂肪の各種類ごとに、ドーナツを揚げたときの吸収量が一連の $4 \times 6 = 24$ 個のデータとして与えられている。このときの種類(あるいは条件)を水準という。いまの場合、水準数 $\alpha=4$ であって、各水準の内部では実験条件は同じであるが、水準の間では効果は異なっているとするのが1元配置である。

なお、ここで種類に加えて例えば揚げる温度を考えるなら、2元配置となるが、これは第3章「実験データの分析」であつかう。

番目の要素は、

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

を表している。

1元配置の例として、4種類の脂肪(ヘッド)を用いたドーナツを揚げたときの、脂肪の吸収量(g、ただし 100 g を減じた数字)のデータをあげておく(表2.2)。いずれも生ドーナツ 1袋当たりの換算値である。なお、原データでは明らかでないが、1度に揚げたものでないこと、全体 24 個がランダムに実験されていることを仮定しておく。各脂肪の種類ごとに一度に揚げたものだと1元配置とはならない。

例2.3 <2次多項式モデル> (2.8) も

$$(2.12) \quad y = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \varepsilon$$

のように(2.9)の形に表せる。 x のデータの値が、行列 X を作っている。

一例として、表2.3のデータはある溶液の成分(I-8)の比率とその溶液の疊り点(透明な溶液が温度の変化によって疊りを生じる温度)の関係を示している。

表2.3' 2次多項式のデータ：

液体のある成分と疊り点の関係

I-8(%)	疊り点
x	y
0	22.1
1	24.5
2	26.0
3	26.8
4	28.2
5	28.9
6	30.0
7	30.4
8	31.4
0	21.9
2	26.1
4	28.5
6	30.3
8	31.5
10	33.1
0	22.8
3	27.3
6	29.8
9	31.8

(出典: ドレーバースミス)

データからは、2量はともに増加(減少)するということしか観察されないが、プロット(図2.4)してみると、あきらかに直線よりは曲線、ことに、2次曲線がよくあてはまることがうかがわれる。これにふさわしいのが、2次多項式のモデルである。

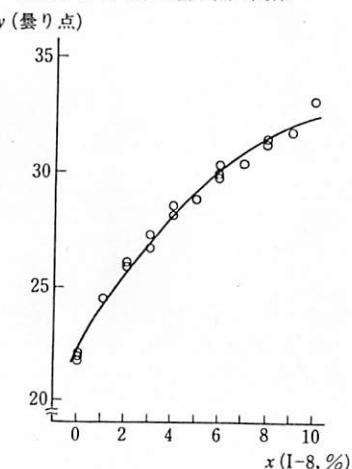
図2.4 2次多項式モデルの例：液体のあ
る成分(I-8, %)と疊り点の関係

表2.3のデータをプロットした点は全体として上方にふくらんでおり、2次多項式があてはまることが予想される。実際に、2.2.2に述べるように、2次多項式

$$y = 22.561 + 1.6680x - 0.06496x^2$$

が、最小二乗法から求められる。全体としてよくあつているが、これはむしろ当初より予想されていることであろう。

I-8 の % を x 、疊り点を y として、プロットすると、プロットした点は全体として上方にふくらんでおり、2次多項式があてはまることが予想される。

このように分散分析や回帰分析は線形モデルを基本としているので、個々の具体的なモデルに替って、一般的な線形モデル(2.9)を考えることにより見通しのよい議論ができるのである。

誤差 ε についての仮定(2.5)は零ベクトル θ 、単位行列 I を用いて

$$(2.13) \quad i) \text{ 不偏性: } E(\varepsilon) = \theta$$

$$(2.14) \quad ii), iii) \text{ 等分散、無相関性: } V(\varepsilon) = \sigma^2 I$$

のように表される。ただし、 $E(\epsilon)$ は平均 $E(\epsilon_i), i=1, \dots, n$ から成る平均ベクトル、同じく $V(\epsilon)$ は共分散 $\text{Cov}(\epsilon_i, \epsilon_j)$ 、分散 $V(\epsilon_i^2) (i=j \text{ のとき})$ から成る分散共分散行列を表す。なお、 $0_n, I_n$ についても次元を表す添字は自明な場合には省略する。

2.2 最小二乗法

2.2.1 最小二乗法の原理

線形モデル(2.9)で、未知母数 θ のある係数 $\mathbf{l} = (l_1, \dots, l_p)'$ による線形結合

$$\mathbf{l}'\theta = l_1\theta_1 + l_2\theta_2 + \dots + l_p\theta_p$$

の線形推定量を考える。たとえば、多項式モデル(2.8)における母数 $\beta_0, \beta_1, \beta_2$ そのもの、あるいは説明変数の値 x における予測値 $\beta_0 + \beta_1x + \beta_2x^2$ などが $\mathbf{l}'\theta$ の例である。これらは $\theta = (\beta_0, \beta_1, \beta_2)'$ に対し、それぞれ

$$\mathbf{l} = (1, 0, 0)', \quad \mathbf{l} = (0, 1, 0)', \quad \mathbf{l} = (0, 0, 1)'$$

あるいは $\mathbf{l} = (1, x, x^2)'$ として $\mathbf{l}'\theta$ の形に表されるからである。

さて、一般の X に対し、 \mathbf{l} が与えられるたびに $\mathbf{l}'\theta$ の BLUE を 2.1.4 の方法で求めるのは可能ではあるにしても煩雑である。そこで \mathbf{l} とは無関係にデータ \mathbf{y} とその期待値の偏差二乗和

$$(2.15) \quad \begin{aligned} S(\theta) &= \|\mathbf{y} - X\theta\|^2 \\ &= (\mathbf{y} - X\theta)'(\mathbf{y} - X\theta), \end{aligned}$$

を最小にする解 $\theta = \hat{\theta}$ を求めておき、単に $\mathbf{l}'\hat{\theta}$ とすることによって $\mathbf{l}'\theta$ の BLUE を求めようというのが最小二乗法の原理 principle of the least square である。

$S(\theta)$ は具体的に、たとえば

$$(2.16) \quad \text{繰り返し測定のモデル(2.2)の場合: } S(\theta) = \sum(y_i - \mu)^2$$

$$(2.17) \quad \text{1元配置モデル(2.7)の場合: } S(\theta) = \sum \sum (y_{ij} - \mu - a_i)^2$$

$$(2.18) \quad \text{多項式モデル(2.8)の場合: } S(\theta) = \sum (y_i - \beta_0 - \beta_1x_i - \beta_2x_i^2)^2$$

のようになる。

これはガウスとマルコフによって証明されたきわめて有用な原理であり、ガウス-マルコフの定理とよばれる。たとえば、(2.16)を μ で微分して 0 とおく

2.2 最小二乗法

と

$$\frac{\partial}{\partial \mu} \sum (y_i - \mu)^2 = -2 \sum (y_i - \mu) = 0$$

より

$$\hat{\mu} = \sum y_i / n = \bar{y}$$

となって、確かに μ の BLUE が得られている。ガウス-マルコフの定理の証明は 2.2.3 で与える。

2.2.2 正規方程式

定理を証明する前に最小二乗法の解 $\hat{\theta}$ の満たす方程式を導いておこう。見当をつけるため、例から入る。

例 2.3' 〈多項式モデル〉 偏差二乗和(2.18)を $\beta_0, \beta_1, \beta_2$ でそれぞれ偏微分して 0 とおいた式から

$$(2.19) \quad \begin{aligned} \partial S / \partial \beta_0 &= -2 \sum (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) \\ &= -2(\sum 1 \cdot y_i - \sum 1 \cdot 1 \beta_0 - \sum 1 \cdot x_i \beta_1 - \sum 1 \cdot x_i^2 \beta_2) = 0 \end{aligned}$$

$$(2.20) \quad \begin{aligned} \partial S / \partial \beta_1 &= -2 \sum x_i (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) \\ &= -2(\sum x_i \cdot y_i - \sum x_i \cdot 1 \beta_0 - \sum x_i \cdot x_i \beta_1 - \sum x_i \cdot x_i^2 \beta_2) = 0 \end{aligned}$$

$$(2.21) \quad \begin{aligned} \partial S / \partial \beta_2 &= -2 \sum x_i^2 (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2) \\ &= -2(\sum x_i^2 \cdot y_i - \sum x_i^2 \cdot 1 \beta_0 - \sum x_i^2 \cdot x_i \beta_1 - \sum x_i^2 \cdot x_i^2 \beta_2) = 0 \end{aligned}$$

を得るが、これらは線形モデル(2.12)の係数行列 X を用いて簡潔に

$$(2.22) \quad X'X\theta = X'y$$

と表すことができる。たとえば、 $X'X$ の(1, 1), (2, 2), (3, 3)要素は、 X の各列の内積で

$$n, \quad \sum x_i^2, \quad \sum x_i^4,$$

となり、(1, 2)及び(2, 1), (1, 3)及び(3, 1), (2, 3)及び(3, 2)要素は、同じく
 $\sum x_i, \quad \sum x_i^2, \quad \sum x_i^3,$

さらに $X'y$ の各要素は、 X の各列と y の内積で

$$\sum y_i, \quad \sum x_i y_i, \quad \sum x_i^2 y_i$$

となることは、行列演算からすぐに確かめられる。

一般の線形モデル(2.9)に対しても(2.22)と同じ結果

(2.23)

$$X'X\theta = X'y$$

が得られる。 (2.23) を θ の正規方程式 normal equation という。この解は推定量を表す意味で $\hat{\theta}$ のようにハットを付けて表す習慣がある。これを満たす $\theta = \hat{\theta}$ が確かに偏差二乗和 $S(\theta)$ を最小にすることは次のように確かめられる。

$$\begin{aligned} S(\theta) &= \{y - X\hat{\theta} + X(\hat{\theta} - \theta)\}'\{y - X\hat{\theta} + X(\hat{\theta} - \theta)\} \\ &= (y - X\hat{\theta})'(y - X\hat{\theta}) + (X\hat{\theta} - X\theta)'(X\hat{\theta} - X\theta) \geq S(\hat{\theta}) \end{aligned}$$

ここで、(2.23) により交叉項(クロス・ターム)は 0 となっている。

一例として、例 2.3 で導入した 2 次多項式が適切と思われる疊り点のデータ ($n=19$) に対して、 $X'X$, $X'y$ を計算してみると(2.23) は 3 元連立 1 次方程式

$$\begin{pmatrix} 19.0 & 84.0 & 550.0 \\ 84.0 & 550.0 & 4068.0 \\ 550.0 & 4068.0 & 32374.0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 531.4 \\ 2536.1 \\ 16994.1 \end{pmatrix}$$

となる。左辺の行列が $X'X$ 、右辺のベクトルが $X'y$ である。

これを満たす $\hat{\theta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)'$ は、左辺の係数の行列 $X'X$ の逆行列を見出すことにより求められる。この例の場合

$$(X'X)^{-1} = \begin{pmatrix} 0.2533559062418 & -0.0971473486910 & 0.0079029364935 \\ 0.0971473486910 & 0.0630038091783 & -0.0062664006288 \\ 0.0079029364935 & -0.0062664006288 & 0.0006840397444 \end{pmatrix}$$

であるから、最小二乗法の解は

$$\hat{\theta} = (X'X)^{-1} \cdot X'y = \begin{pmatrix} 22.5612306265998 \\ 1.6680204365439 \\ -0.0679583610456 \end{pmatrix}$$

となって、求める 2 次多項式は

$$y = 22.561 + 1.6680x - 0.06796x^2$$

であることがわかる。

なお、 x, x^2 のかわりに $x - \bar{x}, (x - \bar{x})^2$ を変数として用いた方が、計算上は楽である。このときは、 $x'x$ の第 1 行、第 1 列は(1, 1)要素以外は 0 となり、3 元連立 1 次方程式は実質上 2 元となるからである(例 2.6 参照)。

ただし、ここで注意しなければいけないのは任意の線形式 $l'\theta$ が推定可能

とは限らないことである。実際 X のランク^{*)}(階数)が未知母数の次元より小さいと(2.23) は不定解の場合に当たり、 $\hat{\theta}$ は一意には定まらない。たとえば 1 元配置のモデル(2.7)では μ と α_i は常に $\mu + \alpha_i$ という結合した形であらわれており、線形な操作だけでは決して μ と α_i は分離できない。

したがって、何か母数に関する付加的な制約(たとえば $\sum \alpha_i = 0$ というよう)を課さない限り μ, α_i それぞれを別個に推定することはできないはずである。実際、この場合の正規方程式はいわゆるランク落ちしており、 $\hat{\mu}, \hat{\alpha}_i$ は不定解となる。そこで、推論の対象を以下に定義する推定可能関数に限る必要がある。不能解の場合でないことは正規方程式(2.23)の右辺が左辺係数行列の列の線形結合で表されることから明らかである(クラーメルの定理)。

推定可能関数　線形モデル(2.9)の母数に関する線形式 $l'\theta$ で、 y の線形式から成る不偏推定量が存在するものを推定可能関数 estimable function という。

定義より明らかに $X\theta$ の各要素およびその線形結合全体が推定可能関数の全体である。なぜなら X およびその線形結合には y および対応する線形結合が不偏推定量となり、逆に y の任意の線形結合の期待値は必ず $X\theta$ の線形結合になるからである。たとえば、繰り返し測定のモデル(2.2)の μ 、1 元配置のモデル(2.7)の $\mu + \alpha_i$ 、 $\alpha_i - \alpha_j$ 、多項式モデル(2.8)の係数 $\beta_0, \beta_1, \beta_2$ などは推定可能関数である。ただし 2 次多項式モデル(2.8)においては少なくとも異なる 3 点以上の x について実験が行われていることが必要である。これに対し、1 元配置の α_i, μ は単独では推定可能ではない。

例 2.4 〈1 元配置モデル〉　データ例 2.2 においても、脂肪 1, 2, 3, 4 の効果 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ を 0 点からはかって

$$72, \quad 85, \quad 76, \quad 62$$

と考えるか、それとも、総平均 73.75 からの各偏差

$$-1.75, \quad 11.25, \quad 2.25, \quad -11.75$$

と考えるかの任意性があり、一通りの決められた推定法があるわけではない。すなわち、各 α_i から μ へ一定量‘しわ寄せ’してもモデルとしては同じものを表すので、 μ, α_i を別個に推定できないのである。一方、どちらに決めて

*) ランク(階数)は章末の付節の 1. でわかりやすく解説してある。

$$\hat{\mu} + \hat{\alpha}_1 = 0 + 72 = 73.75 - 1.75, \quad \hat{\alpha}_1 - \hat{\alpha}_2 = 72 - 85 = -1.75 - 11.25$$

のように推定可能関数に対しては同じ結果が得られる。このことを精密に述べるのが次のガウス-マルコフの定理である。

2.2.3 ガウス-マルコフの定理*

推論を推定可能関数に限ると、次のような最小二乗法の基本定理が成り立つ。

定理 線形モデル(2.9)に関する任意の推定可能関数 $\mathbf{l}'\boldsymbol{\theta}$ について $\mathbf{l}'\hat{\boldsymbol{\theta}}$ が一意に BLUE を与える。ただし、 $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ は(2.23)の正規方程式

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\theta}} = \mathbf{X}'\mathbf{y}$$

を満たす任意の最小二乗解である。これをガウス・マルコフの定理 Gauss-Markov's theorem という。

[証] まず、 x がフル・ランク(つまり、 $\text{rank}(X)$ が $\boldsymbol{\theta}$ の次元 p に等しい)場合の証明はきわめて簡単である。この場合、任意の線形関数が、したがって $\boldsymbol{\theta}$ 自身も、推定可能であり、正規方程式も一意な最小二乗解

$$(2.24) \quad \hat{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

を持つ。

いま任意の線形式 $\mathbf{l}'\boldsymbol{\theta}$ に対し

$$\mathbf{l}'\hat{\boldsymbol{\theta}} = \mathbf{l}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

を構成すると、期待値操作は線形だから

$$E(\mathbf{l}'\hat{\boldsymbol{\theta}}) = \mathbf{l}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{y}) = \mathbf{l}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\theta} = \mathbf{l}'\boldsymbol{\theta}$$

となって $\mathbf{l}'\hat{\boldsymbol{\theta}}$ が $\mathbf{l}'\boldsymbol{\theta}$ の不偏推定量であることがわかる。とくに

$$\mathbf{l} = (0, 0, \dots, \overset{(i)}{1}, 0, \dots, 0)'$$

とおくと

$$E(\hat{\theta}_i) = \theta_i$$

となることに注意しよう。

次に $\mathbf{l}'\hat{\boldsymbol{\theta}}$ とは別に、勝手な線形不偏推定量 $t(\mathbf{y}) = \mathbf{L}'\mathbf{y}$ を考え、両者の差を

$$\mathbf{l}'\hat{\boldsymbol{\theta}} - \mathbf{L}'\mathbf{y} = \{\mathbf{l}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{L}'\}\mathbf{y} \equiv \mathbf{b}'\mathbf{y} \quad (\text{定義})$$

とおく。このとき、不偏性からすべての $\boldsymbol{\theta}$ に対し $E(\mathbf{L}'\mathbf{y}) = \mathbf{l}'\boldsymbol{\theta}$ となるべきで

*) 定理の簡単な場合が本シリーズ I『統計学入門』第13章で扱われている。本節は一般的場合についての厳密な証明である。

あり、よって

$$E(\mathbf{b}'\mathbf{y}) = \mathbf{b}'\mathbf{X}\boldsymbol{\theta} \equiv 0$$

が任意の $\boldsymbol{\theta}$ について成り立たねばならないので

$$(2.25) \quad \mathbf{b}'\mathbf{X} = \mathbf{0}$$

が成り立つ。

次に、一般に 2 つの確率変数 X, Y に対し

$$(2.26) \quad V(X+Y) = V(X) + 2 \text{Cov}(X, Y) + V(Y)$$

が成り立つから、 $\mathbf{L}'\mathbf{y}$ の分散は

$$\begin{aligned} V(\mathbf{L}'\mathbf{y}) &= V(\mathbf{l}'\hat{\boldsymbol{\theta}} - \mathbf{b}'\mathbf{y}) \\ &= V(\mathbf{l}'\hat{\boldsymbol{\theta}}) - 2 \text{Cov}(\mathbf{l}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \mathbf{b}'\mathbf{y}) + V(\mathbf{b}'\mathbf{y}) \end{aligned}$$

と計算できる。一方、誤差 ε が無相関、等分散の仮定(2.14)を満たすとき、2 つの確率変数

$$\mathbf{a}'\mathbf{y} = \sum a_i y_i, \quad \mathbf{b}'\mathbf{y} = \sum b_i y_i$$

※女子が男子をおさえる：競走種目の記録の伸び※

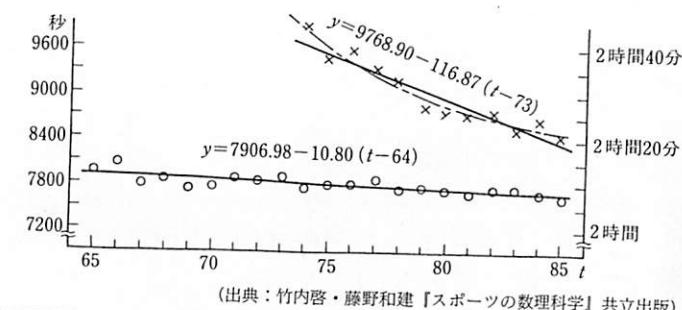
スポーツの記録が向上してきた様子を調べるために、マラソンに対して横軸に年度(西暦末尾 2 術) t 、縦軸にその年度の最高記録をとて平面上にプロットしてみよう。

これに最小二乗法で関係式を推定すると

$$y = 7906.98 - 10.80(t-64) \quad (\text{女子})$$

$$y = 9768.90 - 116.87(t-73) \quad (\text{女子})$$

となって、記録の伸びでは 1 年当たり男子は 11 秒だが、女子は何と 117 秒である。それどころか、主要な競走種目のすべてで、記録の伸びで女子が男子を上まわるという。



の共分散は

$$(2.27) \quad \text{Cov}(a'y, b'y) = \text{Cov}(a'\epsilon, b'\epsilon) = \sum a_i b_i \sigma^2 = (a'b)\sigma^2$$

となることから

$$\text{Cov}(l'(X'X)^{-1}x'y, b'y) = l'(X'X)^{-1}X'b \cdot \sigma^2$$

となる。 (2.25) によりこれは 0 である。よって

$$V(L'y) = V(l'\hat{\theta}) + V(b'y)$$

が成り立つ。ところが、分散は非負なのでこれは

$$V(L'y) \geq V(l'\hat{\theta})$$

を意味する。 $L'y$ は任意の線形不偏推定量だったから、これで $l'\hat{\theta}$ が線形最小分散不偏推定量 BLUE であることが証明された。なお、等号は $b'y=0$ 、すなわち $L'y=l'\hat{\theta}$ のときに成り立つ。(証明終)

次に、 X がランク落ちしている場合には普通の意味では逆行列 $(X'X)^{-1}$ が存在しない。

そこで若干の工夫を要する。その場合、正規方程式(2.23)の任意の不定解を

$$\hat{\theta} = (X'X)^{-1}X'y$$

のように表し、 $(X'X)^{-1}$ の代数的な性質を使って証明する方法や、データを直交変換により扱いやすい標準形に変換して証明するなどの方法がある。ここでは紙数が足りないので省略する。なお、 $(X'X)^{-1}$ は $X'X$ の一般化逆行列とよばれ、統計理論から生れた興味ある線形代数の話題の一つである。

正射影 ここではきわめて直観的な証明を試みよう。最小二乗法の原理(2.15)に立ち戻れば、最小二乗解は $X\theta$ と表されるベクトルの中で観測ベクトル y にユークリッド距離の意味で最も近いベクトルを探すことにはかならない。そこで θ をいろいろ動かしたときにベクトル $X\theta$ の張る空間全体、つまり X の列ベクトル x_1, x_2, \dots, x_p の張る空間

$$\mathcal{L}(X) = \{\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_p x_p ; \quad \theta_1, \theta_2, \dots, \theta_p \text{ は任意の実数}\}$$

を推定空間 $\mathcal{L}(X)$ とよぶことにすると、 $X\hat{\theta}$ は y の $\mathcal{L}(X)$ への正射影 orthogonal projection となるはずである^{*)}(図2.5参照)。したがって、 $X\hat{\theta}$ は一意に定まる。

*) 正射影、射影子の概念に不慣れな読者は章末の付節2を参照。

2.2 最小二乗法

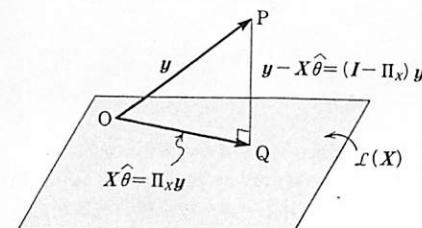


図 2.5 最小二乗法の原理の幾何学的説明

$\mathcal{L}(X)$ は推定の候補となる集合である。その中で、データの点 P から最も近い点 Q は P からの垂線の足、つまり影を落とした点(正射影)であり、その Q を与えるのが $X\hat{\theta}$ である。

なお、 $P \in \mathcal{L}(X)$ なら $Q \equiv P$ となって、正射影が $\mathcal{L}(X)$ の点は動かない。

X がランク落ちしているときはベクトル空間 $\mathcal{L}(X)$ の次元が p より小さく、母数に無駄のある状況であって $\hat{\theta}$ は一意に定まらないがそのことは問題にならない。たとえばモデルを変えることなく無駄な母数を消去すればフル・ランクの場合に帰着し、 $X\hat{\theta}$ の各要素、およびその線形結合が対応する θ の線形式の最小分散不偏推定量であることがいえる。

実際、フル・ランクの場合、 y の $\mathcal{L}(X)$ への正射影

$$X\hat{\theta} = X(X'X)^{-1}X'y$$

を与える行列は

$$(2.28) \quad \Pi_X = X(X'X)^{-1}X'$$

である。この行列はいろいろと有用な性質を持っているので、統計理論ではよく用いられる。 Π_X は対称で、 $\Pi_X^2 = \Pi_X$ (べき等性 idempotency)を満たすから射影子である。また、 Π_X は

$$\Pi_X X\theta = X(X'X)^{-1}X'X\theta = X\theta$$

となって $X\theta$ を動かさないから、 $\mathcal{L}(X)$ への射影子であることがわかる。以後、 Π_X を $\mathcal{L}(X)$ への射影子という。

例 2.5 〈繰返し測定のモデル〉 例 2.1 で扱った一般線形モデル

$$\begin{pmatrix} y \\ (n \times 1) \end{pmatrix} = j\mu + \begin{pmatrix} \epsilon \\ (n \times 1) \end{pmatrix}$$

では推定空間 $\mathcal{L}(j)$ は $j=(1, \dots, 1)'$ の c (任意定数)倍である。 μ の最小二乗推定量は、容易に確かめられるよう

$$\hat{\mu} = (j'j)^{-1}j'y = n^{-1}\sum y_i = \bar{y}$$

であり

$$(2.29) \quad j\hat{\mu} = j\bar{y} = j(j'j)^{-1}j'y$$

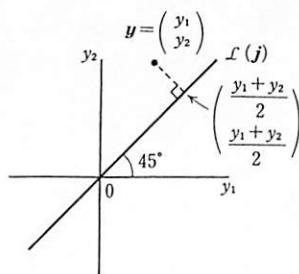


図 2.6 繰返し測定モデルと最小二乗法
部分空間への射影を $n=2$ の場合で見る。
点 (y_1, y_2) から直線 $y_1 = y_2$ への垂線の足をあ
らわし、図 2.5 の特別な場合であるが、その
意味をわかり易い形であらわしている。

において \mathbf{y} の係数行列

$$\Pi_j = j(j'j)^{-1}j' = \frac{1}{n}jj'$$

は対称かつべき等で

$$\Pi_j j \mu = \frac{1}{n} j (j' j) \mu = j \mu$$

のように $\mathcal{L}(j)$ の要素を動かさないから、確かに $\mathcal{L}(j)$ への射影子である。すなわち(2.29)において $j\hat{\mu}$ はデータ \mathbf{y} を推定空間 $\mathcal{L}(j)$ へ正射影したものにほ
かならない。このことをより直観的に理解するために標本数 n が 2 の場合を
図示してみる。その場合、 \mathbf{y} が 2 次元、 $\mathcal{L}(j)$ が 1 次元だから図 2.5 をより具
体的に図 2.6 のような 2 次元平面で表すことができる。

データ \mathbf{y} は 2 次元平面上の 1 点で与えられる。一方、 $\mathcal{L}(j)$ は第 1 座標と第
2 座標の等しい点の全体だから原点を通り 45° の傾きをもった直線である。す
なわち、期待値ベクトル $E(\mathbf{y})$ がこの直線上にあるというのが繰返し測定のモ
デルにほかならない。そして

$$j\hat{\mu} = \begin{pmatrix} \frac{y_1 + y_2}{2} \\ \frac{y_1 + y_2}{2} \end{pmatrix}$$

は、この直線上でデータ \mathbf{y} に最も近い点、すなわち \mathbf{y} の直線 $\mathcal{L}(j)$ への正射
影にほかならない。

例 2.6 <単回帰モデル> よく仮定される関係式

$$(2.30) \quad y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i=1, \dots, n \quad (n \geq 2)$$

で、説明変数 x に対する予測式 $\beta_0 + \beta_1 x$ の BLUE を最小二乗法で求めよう。

モデル(2.30)を

$$(2.30') \quad \begin{aligned} y_i &= \beta_0 + \beta_1 \bar{x} + \beta_1(x_i - \bar{x}) + \varepsilon_i \\ &= \beta_0^* + \beta_1 x_i^* + \varepsilon_i, \quad (\beta_0 \equiv \beta_0^* - \beta_1 \bar{x}) \end{aligned}$$

と書き直しても線形モデルとしてはまったく同じものを表す(推定空間が変わ
らない)。そこで(2.30')の表現で考えると、 $\sum x_i^* = \sum (x_i - \bar{x}) = 0$ であること
から、正規方程式(2.23)はきわめて簡単に

$$\begin{pmatrix} n & 0 \\ 0 & \sum (x_i - \bar{x})^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0^* \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum (x_i - \bar{x}) y_i \end{pmatrix}$$

となる。左辺の行列が $X'X$ だが、その逆行列は簡単に求められる。いま

$$(2.31) \quad S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \bar{x}^2 / n$$

$$(2.32) \quad \begin{aligned} S_{xy} &= \sum (x_i - \bar{x})(y_i - \bar{y}) \\ &= \sum x_i y_i - \bar{x} \cdot \bar{y} / n \end{aligned}$$

を導入すると

$$(X'X)^{-1} = \begin{pmatrix} n^{-1} & 0 \\ 0 & S_{xx}^{-1} \end{pmatrix}$$

から

$$(2.33) \quad \hat{\beta}_0^* = \bar{y}, \quad \hat{\beta}_1 = S_{xy}/S_{xx}$$

が得られる。よって BLUE として

$$(2.34) \quad \hat{\beta}_0 + \hat{\beta}_1 x = \hat{\beta}_0^* + \hat{\beta}_1 x^* = \hat{\beta}_0^* + \hat{\beta}_1 x^* = \bar{y} + (S_{xy}/S_{xx})(x - \bar{x})$$

が得られる。得られた(2.34)は、よく知られているように、データの中心 (\bar{x}, \bar{y}) を通り、傾き

$$\frac{S_{xy}}{S_{xx}} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

の直線を表す。

(2.30)のまま正規方程式を立てても同じ結果が得られるが、逆行列の計算がやや複雑になる。

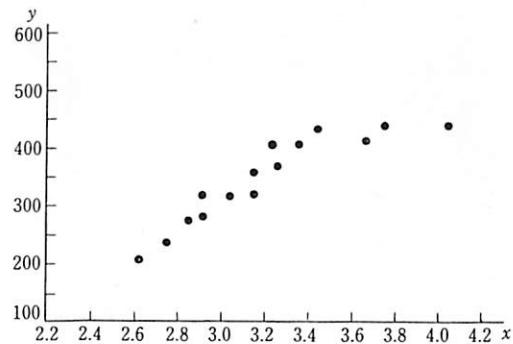
例 2.6' <単回帰モデル> 表 2.4 は対数積算温度 x (コンクリート養生中

表 2.4^t 単回帰のデータ：コンクリートの養生の温度条件と圧縮強度の関係
対数積算温度(x)とコンクリート圧縮強度(y)の関係

データ番号	対数積算温度	圧縮強度
1	2.62	209
2	2.75	237
3	2.85	273
4	2.92	316
5	2.92	283
6	3.05	317
7	3.14	321
8	3.15	356
9	3.23	404
10	3.26	368
11	3.36	405
12	3.44	429
13	3.66	413
14	3.74	440
15	4.04	439

(出典：広津(編) (1993))

図 2.7, 2.8 を参照。分析については 2.3 節を見よ。



(出典：広津(編) (1993))

図 2.7 単回帰のモデルの例：コンクリートの養生の温度条件と圧縮強度の関係(表 2.4 参照)
直線をあてはめたが、なお若干の議論の余地がある。

の温度と時間の積の対数)とそれに対して得られたコンクリート圧縮強度 y を 15 通り測定したものである。データ 1~15 を散布図に描くと図 2.7 のようになる。

2.2 最小二乗法

一見して、2 次多項式があてはまりそうであるが、それは最右端のデータ点 (#15) がややとび離れているためとも思われる。そこで、一応は、直線による単回帰のモデルをあてはめた(#15 を除くと、直線性がより濃厚になるもの、それでも 2 次多項式の可能性は捨てられない)。

このデータに対しては、(2.31), (2.32) の後の式から次のように計算される。

$$S_{xx} = 2.1419, \quad S_{xy} = 307.3467$$

例 2.7 <1 元配置モデル> (2.11) の場合、母数ベクトルの次元 $p=a+1$ に対し $\text{rank}(X)=a$ でランク落ちしている。実際、正規方程式は

$$(2.35) \quad \begin{pmatrix} ar & r \cdots r \\ r & r \cdots r \\ \vdots & \ddots \\ r & r \end{pmatrix} \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_a \end{pmatrix} = \begin{pmatrix} y_{..} \\ y_{1..} \\ \vdots \\ y_{a..} \end{pmatrix},$$

$$y_{i..} = \sum_{j=1}^r y_{ij}, \quad y_{..} = \sum_{i=1}^a \sum_{j=1}^r y_{ij}$$

の形となる。この係数行列がランク落ちしたため $\hat{\mu}, \hat{\alpha}_1, \dots, \hat{\alpha}_a$ は一意には定まらない。たとえば、水準数が $a=2$ のとき、(2.35) の係数行列の行列式は

$$\begin{vmatrix} 2r & r & r \\ r & r & 0 \\ r & 0 & r \end{vmatrix} = 0$$

となって、(2.35) は一意に解けないのである。

しかしながら推定可能関数については方程式(2.35)から一意に推定量が定まるはずである。たとえば $\mu + \alpha_i$ については(2.35) の第 $i+1$ 行目から

$$\widehat{\mu + \alpha_i} = \bar{y}_{i..}$$

が得られ、それから

$$(2.36) \quad \widehat{\alpha_i - \alpha_j} = \widehat{\mu + \alpha_i} - \widehat{\mu + \alpha_j} = \bar{y}_{i..} - \bar{y}_{j..}$$

が得られる。これらはガウス-マルコフの定理により、一意な最小二乗推定量であり、BLUE を与える。

2.2.4 最小二乗法と制約式^{*}

1 元配置のモデル(2.7)あるいは(2.11)で $\mu=0$ と置いても、単に第 i 処理の

*) 初読の際はとばしてよい。

平均を α_i と表しただけで線形モデルとしては何も変わらず、推定可能関数の全体にも変化がない。しかるに X のランク落ちは解消し、一意に $\hat{\alpha}_i = \bar{y}_{i..}$ 、
処理差 $\hat{\alpha}_i - \hat{\alpha}_{i'} = \bar{y}_{i..} - \bar{y}_{i'..}$ が求まる。

一方、それは正規方程式(2.35)に

$$(2.37) \quad \mu = 0$$

という制約式を付加して解いたものとも一致する。ちなみに(2.35)と(2.37)を連立させて解くと一意に

$$(2.38) \quad \hat{\mu} = 0, \quad \hat{\alpha}_i = \bar{y}_{i..}$$

が得られ、これらを合成すれば、(2.36)と同じ推定可能関数の推定値

$$\hat{\mu} + \hat{\alpha}_i = \bar{y}_{i..}, \quad \hat{\alpha}_i - \hat{\alpha}_{i'} = \bar{y}_{i..} - \bar{y}_{i'..}$$

が求まる。

たとえば、2.2.2 でもふれたように1元配置モデルのデータ例2.1において、脂肪1, 2, 3, 4 の効果 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ を

$$72, \quad 85, \quad 76, \quad 62$$

と推定するのは $\mu=0$ と仮定した場合であり

$$-1.75, \quad 11.25, \quad 2.25, \quad -11.75$$

という推定値は、 $\sum \alpha_i = 0$ という制約から求められる。どう考えるかは基本的には任意だから、他の制約(たとえば $\alpha_1 = 0$)も可能であるが、いずれにせよ、何らかの制約なしには、 α_i および μ は推定することができない。ただし、この制約はどんなものでもよいというわけではない。

正規方程式が一意に解け、しかも推定可能関数には影響を与えない制約式 $H\theta = 0$ の条件は次の(1), (2)である*。

(1) X の行ベクトルの張る空間と H の行ベクトルの張る空間は $\mathbf{0}'$ 以外に共通部分を持たない。

$$(2) \quad \text{rank} \begin{pmatrix} X \\ H \end{pmatrix} = p$$

(1)は要するに $H\theta = 0$ が推定可能関数を0とおくようなことになつてゐないため、(2)は正規方程式に $H\theta = 0$ を付け加えることにより解が一意に定まるための条件である。

制約式 $H\theta = 0$ を付け加えることはモデル(2.9)の実質的なパラメータを減らす(一部消去する)ことに相当する。そして上記の条件(1), (2)は推定空間 $L(X)$ を変えず

に、フル・ランクの表現を得るために条件ということもできる。そのためこのような制約式を‘母数の無駄を省くための制約式’とよぶ。

例 2.7' <1元配置のモデル> (2.11)では、ふつう

$$(a) \quad \mu = 0, \quad (b) \quad \alpha_1 = 0, \quad (c) \quad \sum \alpha_i = 0$$

のどれかを制約式とする。いずれもが上の条件(1), (2)を満たしている。

(a)について既に述べた通りなので、(b), (c)について検討する。(b)の場合、(2.35)と $\alpha_1 = 0$ から

$$(2.39) \quad \hat{\mu} = \bar{y}_{1..}, \quad \hat{\alpha}_1 = 0, \quad \hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{1..} \quad (i=2, \dots, a)$$

となる。これは当然(2.38)とは異なるが推定可能関数については

$$\hat{\mu} + \hat{\alpha}_i = \bar{y}_{i..}, \quad \hat{\alpha}_i - \hat{\alpha}_j = \bar{y}_{i..} - \bar{y}_{j..}$$

となって(2.36)と同じ結果が得られる。同様に(c)については

$$(2.40) \quad \hat{\mu} = \bar{y} \text{ (総平均)}, \quad \hat{\alpha}_i = \bar{y}_{i..} - \bar{y}$$

となり、これは(2.38)および(2.39)と異なるが、推定可能関数については(2.36)と同じ結果が得られる。

2.3 最小二乗推定量の分散

最小二乗推定量は線形推定量である。一般に線形推定量は

$$t(\mathbf{y}) = L'y = \sum L_i y_i$$

と表すことができる。したがって、その分散は

$$(2.41) \quad V(L'y) = V(\sum L_i y_i) = \sum L_i^2 V(y_i) = L'L \cdot \sigma^2$$

のように求めることができる。とくに X がフル・ランクの場合、推定量

$$l'\hat{\theta} = l'(X'X)^{-1}X'y$$

の分散は

$$(2.42) \quad l'(X'X)^{-1}X' \cdot X(X'X)^{-1}l \cdot \sigma^2 = l'(X'X)^{-1}l \cdot \sigma^2$$

で与えられる。

とくに $\hat{\theta}_i$ 自体の分散は $l = (0, \dots, 0, 1, 0, \dots, 0)'^{(i)}$ とおけば、 $X'X$ の対角線要素から

$$V(\hat{\theta}_i) = (X'X)^{-1}\sigma^2 \text{ の } (i, i) \text{ 要素}$$

として求められる。同様に

$$\text{Cov}(\hat{\theta}_i, \hat{\theta}_j) = (X'X)^{-1}\sigma^2 \text{ の } (i, j) \text{ 要素}$$

のように表すことができる。

例 2.8 <繰返し測定のモデル> (2.10)の場合

$$\hat{\mu} = (j'j)^{-1}j'y = \bar{y}$$

の分散は

$$(j'j)^{-1}\sigma^2 = \sigma^2/n$$

で与えられる。ここで、 σ^2 は(2.5)の誤差の分散である。これはもちろん、 \bar{y} の分散を直接計算して得られる周知のものと一致する。

例 2.9 <単回帰モデル> 回帰分析でよく用いる

$$(2.43) \quad y_i = \beta_0^* + \beta_1(x_i - \bar{x}) + \varepsilon_i$$

の場合、係数行列は

$$(2.44) \quad X = (j \quad x^*), \quad x^* = (x_1 - \bar{x}, \dots, x_n - \bar{x})'$$

という形をしており、既に見たように

$$X'X = \begin{pmatrix} j'j & jX^* \\ X^{*\prime}j & X^{*\prime}X^* \end{pmatrix} = \begin{pmatrix} n & O \\ O & S_{xx} \end{pmatrix}$$

だから

$$(X'X)^{-1} = \begin{pmatrix} n^{-1} & O \\ O & S_{xx}^{-1} \end{pmatrix}$$

である。推定量は(2.33)で与えられるが、 $\theta = (\beta_0^*, \beta_1)', l = (1, 0)', (0, 1)'$ として、(2.42)を用いると、これらについてよく知られた結果

$$(2.45) \quad V(\hat{\beta}_0^*) = \sigma^2/n, \quad V(\hat{\beta}_1) = \sigma^2/S_{xx} = \sigma^2/\sum(x_i - \bar{x})^2$$

が得られる。また

$$\beta_0^* + \beta_1(x - \bar{x}) = (1, x - \bar{x})\theta$$

と表せるから、(2.42)に代入すると、予測式のばらつきとして

$$(2.46) \quad \begin{aligned} V(\hat{\beta}_0^* + \beta_1(x - \bar{x})) &= (1, x - \bar{x}) \begin{pmatrix} n^{-1} & 0 \\ 0 & S_{xx}^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix} \sigma^2 \\ &= \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}} \right) \sigma^2 \end{aligned}$$

が得られる。

2.4 誤差分散 σ^2 の推定

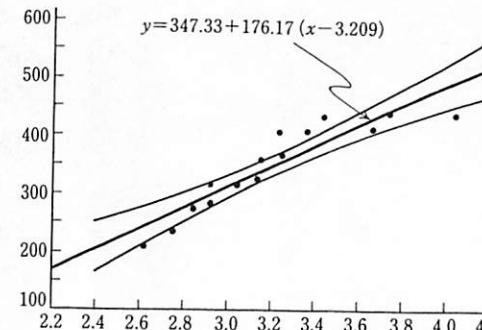


図 2.8 単回帰モデルのあてはめの例：コンクリートの養生の温度条件と圧縮強度の関係および推定量の限界

図 2.7 に対し、(2.46)による標準偏差の 2.16 倍 ($t_{0.025}(13)=2.16$) の限界を入れた。(表 2.4, 図 2.7 参照)

これもよく知られた結果である。(2.46)はデータの中心 $x = \bar{x}$ で最小値 σ^2/n をとり、そこから離れるに従って大きくなる。とくにデータの存在範囲を超えて外挿したような場合は推定量の分散が相当大きくなることに注意しなければならない。

例 2.6' でとり上げた単回帰分析について見てみよう。推定された直線は、

$$\hat{\beta}_0^* = \bar{y} = 347.33, \quad \hat{\beta}_1 = \frac{377.3467}{2.1419} = 176.17$$

から

$$y = 347.33 + 176.17(x - 3.209)$$

であって、図 2.8 のようである。この回帰直線に(2.46)の平方根として与えられる標準偏差によって幅を付した上下の限界が、双曲線

$$\begin{aligned} y &= \hat{\beta}_0^* + \hat{\beta}_1(x - \bar{x}) \pm \sqrt{\left\{ \frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}} \right\} \sigma^2} \\ &= 347.33 + 176.17(x - 3.209) \pm \sqrt{62.055 + 434.58(x - \bar{x})^2} \end{aligned}$$

として示されている。ただし、(2.46)で σ^2 の値は $\sigma^2 = 930.825$ と推定されている(この推定については、例 2.11 で後述する)。

2.4 誤差分散 σ^2 の推定

最小二乗法により線形モデル(2.9)をあてはめた後の残差 $y - X\hat{\theta}$ の二乗和

$$S(\hat{\theta}) = (y - X\hat{\theta})'(y - X\hat{\theta})$$

は残差平方和とよばれ、線形モデル(2.9)の誤差分散 σ^2 に関する情報を担っている。 $X\hat{\theta}=\Pi_Xy$ が y の推定空間 $\mathcal{L}(X)$ への正射影であったのに対し、残差

$$y-X\hat{\theta}=(I-\Pi_X)y$$

は、2.2.3の正射影の考え方でも見たように、 $\mathcal{L}(x)$ の直交補空間への正射影である(図2.5参照)。

ここで、残差平方和 $S(\hat{\theta})$ の大きさについて、次の結果を示そう。まず、その期待値は、誤差分散 σ^2 (すなわち(2.5))および X のランクによって

$$(2.47) \quad E(S(\hat{\theta}))=(n-q)\cdot\sigma^2, \quad q=\text{rank}(X)$$

とあらわされる。したがって

$$(2.48) \quad \nu_e=n-q$$

とし、それによって

$$(2.49) \quad \begin{aligned} \hat{\sigma}^2 &= S(\hat{\theta})/\nu_e \\ &= (y-X\hat{\theta})'(y-X\hat{\theta})/\nu_e \end{aligned}$$

とおくと、これが誤差分散 σ^2 の不偏推定量を与える。 ν_e は自由度とよばれる。

$S(\hat{\theta})$ の期待値(2.47)の証明は次の通りである。ここでは行列を用いる。

$$(2.50) \quad \begin{aligned} E(S(\hat{\theta})) &= E\{\text{tr}(y-X\hat{\theta})(y-X\hat{\theta})'\} \\ &= E\{\text{tr}((y-X\hat{\theta})(y-X\hat{\theta})')\} \quad (\text{tr } AB = \text{tr } BA) \\ &= \text{tr}\{E((y-X\hat{\theta})(y-X\hat{\theta})')\} \quad (E \text{ の線形性}) \\ &= \text{tr}\{V(y-X\hat{\theta})\} \quad (V \text{ は分散共分散行列}) \\ &= \text{tr}\{V((I-\Pi_X)y)\} \quad (\text{残差の書き換え}) \\ &= \text{tr}\{V(I-\Pi_X)(I-\Pi_X)'\} \quad (\text{後述}) \\ &= (\text{tr}(I-\Pi_X))\cdot\sigma^2 \quad (\text{対称性, べき等性}) \\ &= \text{rank}(I-\Pi_X)\sigma^2 \quad (\text{射影子の性質}) \\ &= (n-\text{rank}(X))\cdot\sigma^2 \quad (\text{射影子の性質 iii'}) \end{aligned}$$

ここで、任意の確率変数のベクトル ε に対し、行列 A による変換が行なわれれば、分散共分散行列は

$$(2.51) \quad V(A\varepsilon)=AV(\varepsilon)A'$$

と変換されること、および ε が(2.5)のように無相関、つまり

$$(2.52) \quad V(\varepsilon)=\sigma^2 I$$

ならば

$$(2.53) \quad V(A\varepsilon)=\sigma^2 AA'$$

となる性質などを、(2.50)に用いた。

残差平方和の計算は次の式が便利である。

$$(2.54) \quad S(\hat{\theta})=(y-X\hat{\theta})'(y-X\hat{\theta})=y'y-\hat{\theta}'\cdot(X'y)$$

(2.54)において第1項はデータの二乗和、第2項は $\hat{\theta}$ と正規方程式の右辺 $X'y$ との内積であり、 n 個の積の計算ですむ。

例2.10 <繰返し測定のモデル> (2.10)の場合、(2.54)に対応して

$$S(\hat{\mu})=\sum(y_i-\hat{\mu})^2=\sum y_i^2-\hat{\mu}\cdot y=\sum y_i^2-y.^2/n$$

となる。ここに $\hat{\mu}=\sum y_i/n=y./n=\bar{y}$ である。自由度は $n-1$ だから

$$(2.55) \quad \hat{\sigma}^2=\left(\sum y_i^2-\frac{y.^2}{n}\right)/(n-1)$$

が得られる。これはその性質からふつう不偏分散 s^2 として

$$s^2=\sum(y_i-\bar{y})^2/(n-1)$$

と定義しているものである。

例2.1 に対しては

$$y.=\sum y_i=339.08, \quad \sum y_i^2=12775.1042, \quad n=9$$

であるから

$$\hat{\sigma}^2=(12775.1042-(339.08)^2/9)/8=0.0096$$

となる。これが偶然誤差の分散の推定値である。

例2.11 <単回帰モデル> (2.30)の場合

$$(2.56) \quad \begin{aligned} S(\hat{\beta}_0^*, \hat{\beta}_1) &= \sum(y_i-\hat{\beta}_0^*-\hat{\beta}_1(x_i-\bar{x}))^2 \\ &= \sum y_i^2-\bar{y}\cdot y-(S_{xy}/S_{xx})\cdot S_{xy} \\ &= \sum y_i^2-y.^2/n-S_{xy}^2/S_{xx} \end{aligned}$$

となる。自由度は $n-2$ だから

$$(2.57) \quad \hat{\sigma}^2=\left(\sum y_i^2-\frac{y.^2}{n}-\frac{S_{xy}^2}{S_{xx}}\right)/(n-2)$$

が不偏分散である。

例2.6' に対して、例2.9における分析を続ける。直前例のように

$$S_{yy}=\sum(y_i-\bar{y})^2=\sum y_i^2-y.^2/n=78,579.3333$$

$$S_{xx}=2.1419, \quad S_{xy}=377.3467$$

と計算すると

$$\begin{aligned}\hat{\sigma}^2 &= \frac{78,579.3333 - (377.3467)^2 / 2.1419}{15-2} \\ &= 930.825\end{aligned}$$

と算出される。図2.8の限界の幅では σ^2 の代りにこれを用いたのである。

※残りものに情報※

残差とは、説明されずに残ったものという意味であり、直線、曲線をあてはめた後、残った何か最後の‘余りもの’という印象を与える。

だが、直線や曲線を引いたあと、もとのデータがこれらに実際に合致しているか、していないとすれば、どこがどのように、どのような原因や理由で（もしわかれれば）そうなるのかをくわしく調べることが重要である。そのような分析は残差分析とよばれ、種々の方法が研究されている。

次の例は、すべて異なった様子のデータなのに、同一の直線があてはまってしまう有名な例である。この辺が統計学の深いところである。なお、データは章末の練習問題2.4に収録されている。

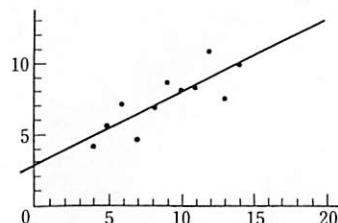


図1 ほどよい例

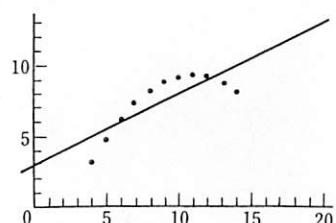


図2 直線をあてはめるべきでなかった例

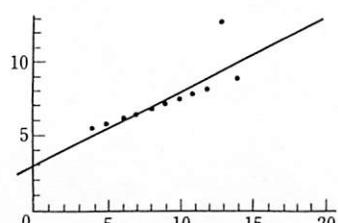


図3 1つを除けば、別の直線が完全にあてはまつた例

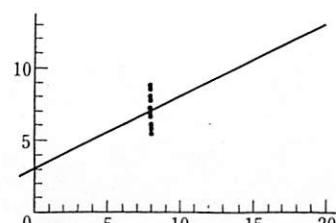


図4 1つを除けば、直線の傾きさえ出なかつた例

(出典: F. J. Anscombe (1973) *American Statistician* 27, pp. 17-21.)

2.5 正規線形モデルと標本分布

前節まで扱った線形モデル(2.9)

$$(2.9') \quad y = X\theta + \varepsilon$$

では、誤差に関し不偏性 $E(\varepsilon) = 0$ 、等分散、無相関性 $V(\varepsilon) = \sigma^2 I$ だけを仮定し、特別な分布は想定していない。この節では誤差分布がこれらの仮定に加え、さらに正規分布であるとしたときの最小二乗推定量の諸性質について述べる。なお、このとき(2.58)を正規線形モデル normal linear model という。一般にデータ y から導かれる統計量の従う分布を標本分布という。この節では(2.58)の誤差がたがいに独立に正規分布 $N(0, \sigma^2)$ に従うときの最小二乗推定量や残差平方和の標本分布を考察しよう。その結果は次節で検定の問題を考えるときに使われる。

2.5.1 最小二乗推定量の標本分布

正規線形モデルの場合、BLUE はさらに強い最適性をもつことが示される。すなわち、それは推定量を y の線形関数に限らず、すべての不偏推定量の中で考えても最小分散の不偏推定量であることがいえる（証明はやや高度なので省略する）。

最小二乗推定量は、 y の線形結合だから、再び正規分布に従うので、平均と分散が求まれば標本分布が定まることになる。任意の推定可能関数 $l'\theta$ についても、その BLUE である最小二乗推定量 $l'\hat{\theta}$ の期待値は $l'\theta$ 、分散は2.3節で述べたように、この $l'\hat{\theta}$ が

$$l'\hat{\theta} = \sum L_i y_i$$

のように表現されているときには

$$V(\sum L_i y_i) = \sum L_i^2 \cdot \sigma^2$$

と表される。

したがって、最小二乗推定量 $l'\hat{\theta}$ は正規分布

$$N(l'\theta, \sum L_i^2 \sigma^2)$$

に従う。なお、 X がフル・ランクの場合には、分散は(2.42)で与えられる*。

* 章末の付節の4にも、簡単な要約を与えてある。

2.5.2 残差平方和の標本分布

線形モデル(2.9)に関して、残差平方和

$$S(\hat{\theta}) = (\mathbf{y} - \mathbf{X}\hat{\theta})'(\mathbf{y} - \mathbf{X}\hat{\theta})$$

を自由度

$$\nu_e = n - q \quad (q = \text{rank } (\mathbf{X}))$$

で除した

$$(2.58) \quad S(\hat{\theta})/\nu_e$$

が σ^2 の不偏推定量であることは既に示した。

ここで誤差に正規性を仮定すると、(2.58)が σ^2 のあらゆる不偏推定量の中で最小分散であることもいえる。

以下では、 σ^2 を推定する残差平方程について次の重要な性質を導こう。すなわち、

$$(2.59) \quad S(\hat{\theta})/\sigma^2$$

の分布は自由度 ν_e の χ^2 分布 $\chi^2(\nu_e)$ に従う。

(2.49)の不偏分散 $\hat{\sigma}^2$ もあらゆる σ^2 の不偏推定量のクラスで最小分散であり、自由度 ν の χ^2 分布 $\chi^2(\nu)$ の分散が 2ν であることから、この $\hat{\sigma}^2$ 分散は

$$(2.60) \quad V(\hat{\sigma}^2) = (\sigma^2/\nu_e)^2 \cdot 2\nu_e = 2\sigma^4/\nu_e$$

で与えられる。さらに

$$(2.61) \quad \text{Cov}(\mathbf{X}\hat{\theta}, \mathbf{y} - \mathbf{X}\hat{\theta}) = \text{Cov}(\Pi_X \mathbf{y}, (\mathbf{I} - \Pi_X)\mathbf{y}) = \sigma^2 \Pi_X (\mathbf{I} - \Pi_X) = 0$$

となることから、 $\hat{\sigma}^2$ は $\mathbf{X}\hat{\theta}$ の各要素と独立であることもわかる。

上に述べたことをもう少し確認しておこう。まず $S(\hat{\theta})$ が \mathbf{y} の行列 $\mathbf{I} - \Pi_X$ に関する2次形式であることを今一度確認しておこう。そして、線形代数学の次の定理を用いる。

一般に $n \times n$ 實対称行列 A については、 λ_i, \mathbf{l}_i を A の固有値、およびそれに対応する固有ベクトルとして

$$(1) \quad A = \sum_{i=1}^n \lambda_i \mathbf{l}_i \mathbf{l}_i'$$

という分解^{*)}(スペクトル分解 spectral resolution といわれる)が成り立つ。ただし、 \mathbf{l}_i は固有ベクトルの直交性をもとに

^{*)} 線形代数学ではよく知られている。なお、特異値分解ともいう。

$$(2) \quad \mathbf{l}_i' \mathbf{l}_j = \delta_{ij}, \quad i, j = 1, \dots, n$$

と規準化されている(δ_{ij} はクロネッカーデルタで、 $i=j$ のとき 1、それ以外では 0 である)。

とくに射影行列 $\mathbf{I} - \Pi_X$ の場合は、固有値は ν_e 個の 1 と $n - \nu_e$ 個の 0 から成る。そこで一般性を失うことなく $\mathbf{l}_1, \dots, \mathbf{l}_{\nu_e}$ を固有値 1 に対する固有ベクトルとして、分解

$$(3) \quad \mathbf{I} - \Pi_X = \sum_{i=1}^{\nu_e} \mathbf{l}_i \mathbf{l}_i'$$

が成り立つ。これを $S(\hat{\theta})$ の定義式に代入すると

$$(4) \quad S(\hat{\theta}) = \sum_{i=1}^{\nu_e} \mathbf{l}_i' \mathbf{y} \mathbf{l}_i \mathbf{l}_i' \mathbf{y} = \sum_{i=1}^{\nu_e} (\mathbf{l}_i' \mathbf{y})^2$$

という表現が得られる。これらの変数(群)

$$\mathbf{l}_1 \mathbf{y}, \quad \mathbf{l}_2 \mathbf{y}, \dots, \mathbf{l}_{\nu_e} \mathbf{y}$$

については、まず固有ベクトルの性質(2)により

$$V(\mathbf{l}_i' \mathbf{y}) = \mathbf{l}_i' V(\mathbf{y}) \mathbf{l}_i = \mathbf{l}_i' (\sigma^2 \mathbf{I}) \mathbf{l}_i = \sigma^2,$$

$$\text{Cov}(\mathbf{l}_i' \mathbf{y}, \mathbf{l}_j' \mathbf{y}) = \mathbf{l}_i' \text{Cov}(\mathbf{y}, \mathbf{y}) \mathbf{l}_j = \mathbf{l}_i' (\sigma^2 \mathbf{I}) \mathbf{l}_j = 0 \quad (i \neq j)$$

が成り立つ。一方

$$(\mathbf{I} - \Pi_X) \mathbf{X} = \mathbf{0}$$

に(3)を代入するとわかるように

$$\mathbf{l}_i' \mathbf{X} = \mathbf{0} \quad (i = 1, \dots, \nu_e)$$

が成り立つので、 $\mathbf{l}_i' \mathbf{y}$ ($i = 1, \dots, \nu_e$) の期待値は

$$E(\mathbf{l}_i' \mathbf{y}) = 0 \quad (i = 1, \dots, \nu_e)$$

である。以上により、 $S(\hat{\theta})$ がたがいに独立に正規分布 $N(0, \sigma^2)$ に従う ν_e 個の正規変数 $\mathbf{l}_i' \mathbf{y}$ ($i = 1, 2, \dots, \nu_e$) の二乗和となることがわかる。

とくに σ^2 で割って規準化すると

(5) $S(\hat{\theta})/\sigma^2$ がたがいに独立な標準正規分布に従う変数 ν_e 個の二乗和が成り立つが、(5)の右辺の分布は自由度 ν_e の χ^2 分布 $\chi^2(\nu_e)$ として知られている。

これで証明は終りであるが、補足として次の性質をつけ加えておこう。自由度 ν の χ^2 分布 $\chi^2(\nu)$ にしたがう確率変数 χ^2 の期待値と分散はそれぞれ

$$E(\chi^2) = \nu, \quad V(\chi^2) = 2\nu$$

である。したがって

$$E(S(\hat{\theta})/\sigma^2) = \nu_e, \quad V(S(\hat{\theta})/\sigma^2) = 2\nu_e$$

が成り立つ。これから

$$\hat{\sigma}^2 = S(\hat{\theta})/\nu_e$$

の期待値、分散が

$$E(\hat{\sigma}^2) = (\sigma^2/\nu_e) \cdot \nu_e = \sigma^2, \quad V(\hat{\sigma}^2) = (\sigma^2/\nu_e)^2 \cdot 2\nu_e = 2\sigma^4/\nu_e$$

となって、 $\hat{\sigma}^2$ が σ^2 の不偏推定量であることおよび分散が(2.61)で与えられることが確かめられる。

さらに、 $\mathbf{l}_i' \mathbf{X} = \mathbf{0}$ ($i = 1, \dots, \nu_e$) であることから $S(\hat{\theta})$ の各成分と推定可能関数の推定量全体 $\mathbf{X}\hat{\theta}$ は無相関であり、正規分布の場合、それは $S(\hat{\theta})$ と $\mathbf{X}\hat{\theta}$ の独立を意味

する。

例 2.12 <繰返し測定のモデル> 例 2.5 で $n=3$ の場合を考える。この場合 $\nu_e = n-1=2$ だから、残差平方和

$$S(\hat{\theta}) = \sum_1^3 (y_i - \bar{y})^2$$

は、 $y=(y_1, y_2, y_3)'$ から導かれ、かつ $N(0, \sigma^2)$ に従うある 2 つの互いに独立な正規確率変数 $l_1 y, l_2 y$ の二乗和で表されるはずである。これを確かめよう。

例 2.5 より、一般に

$$\Pi_j = \frac{1}{n} j j'$$

である。いまは $n=3$ とすると

$$(6) \quad I - \Pi_j = I_3 - \frac{1}{3} j_3 j_3' \\ = \begin{pmatrix} 1 - \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & 1 - \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & 1 - \frac{1}{3} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

の行和はすべて 0 だから固有値 0 に対する自明な固有ベクトル j を持つ。一方、 j に直交する任意のベクトル

(7) $l = (l_1, l_2, l_3)', \quad l' j = l_1 + l_2 + l_3 = 0$
に対して、すぐ確認できるように

$$\frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} l = l$$

が成り立ち、逆もまた真であるから、 l は固有値 1 に対する固有ベクトルである*。そこで(7)を満たし、たがいに直交し、かつ長さ 1 に規準化されたベクトル l_1, l_2 を選べば、それが $S(\hat{\theta})$ の分解(4)を与える y の係数である。そのようなベクトルの組は無数にあり、たとえば

$$(8) \quad l_1 = (1/\sqrt{2}, -1/\sqrt{2}, 0)' \quad l_2 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})'$$

あるいは

$$(9) \quad l_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})' \quad l_2 = (1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})'$$

などを選ぶことができる。読者は、これらを用いた次の 2 つの等式が成り立つことを確認して欲しい。

*) 固有方程式を解かなくてすむ‘うまい’方法である。形どおりの解き方は章末の付節の 3 参照。

$$(10) \quad \sum_1^3 (y_i - \bar{y})^2 = (y_1/\sqrt{2} - y_2/\sqrt{2})^2 + (y_1/\sqrt{6} + y_2/\sqrt{6} - 2y_3/\sqrt{6})^2$$

$$(11) \quad \sum_1^3 (y_i - \bar{y})^2 = (-y_1/\sqrt{2} + y_3/\sqrt{2})^2 + (y_1/\sqrt{6} - 2y_2/\sqrt{6} + y_3/\sqrt{6})^2$$

(10) や (11) で () 内の 2 つの成分(線形結合)はたしかに期待値 0、分散 σ^2 に規準化されており、たがいに独立である。

2.6 線形仮説の検定

2.6.1 t 検定と F 検定

正規線形モデルを規定する母数に関する線形仮説の検定を考える。たとえば、繰返し測定のモデル(2.10)において、真の値は 0 であるという仮説

$$(2.62) \quad H_0: \mu = 0,$$

また、1 元配置のモデル(2.11)における効果の一様性を表す仮説

$$(2.63) \quad H_0: \alpha_1 = \dots = \alpha_a,$$

さらには、2 次回帰モデル(2.12)において、2 次の項がないという仮説

$$(2.64) \quad H_0: \beta_2 = 0$$

などは、すべて一般の線形モデル(2.9)における線形制約

$$(2.65) \quad H_0: K\theta = 0$$

の形に表され、線形仮説 linear hypothesis とよばれる。

なぜなら、たとえば(2.64)は

$$0 \cdot \beta_0 + 0 \cdot \beta_1 + 1 \cdot \beta_2 = 0$$

とあらわされ、(2.63)は

$$0 \cdot \mu + 1 \cdot \alpha_1 - 1 \cdot \alpha_2 + 0 \cdot \alpha_3 + \dots + 0 \cdot \alpha_a = 0$$

$$0 \cdot \mu + 1 \cdot \alpha_1 + 0 \cdot \alpha_2 - 1 \cdot \alpha_3 + \dots + 0 \cdot \alpha_a = 0$$

...

$$0 \cdot \mu + 1 \cdot \alpha_1 + 0 \cdot \alpha_2 + 0 \cdot \alpha_3 + \dots - 1 \cdot \alpha_a = 0$$

とあらわされるからである。

もっとも、1 元配置の場合に(2.63)の代りに

$$(2.66) \quad H_0: \alpha_1 = \dots = \alpha_a = 0$$

としても仮説の下での線形モデルは同じものを表す。しかるに(2.63)と(2.66)とは制約式としては違うものを表すので、帰無仮説は(2.65)のような形式で表すより、その意味する線形モデルで表す方が混乱が少ない。たとえば仮説(2.63), (2.66)はともに

$$(2.67) \quad H_0: E(\mathbf{y}) = j\mu$$

で表される。

t検定 仮説(2.62)や(2.64)のような線形仮説の検定は明解である。これは一般的にいって、ある推定可能関数に関する帰無仮説

$$(2.68) \quad H_0: \mathbf{l}'\boldsymbol{\theta} = 0$$

の検定にほかならない。そこでBLUEを \mathbf{y} の関数として

$$\mathbf{l}'\hat{\boldsymbol{\theta}} = \mathbf{L}'\mathbf{y}$$

と表すと、2.5.1によりその仮説の下での分布は正規分布 $N(0, \mathbf{L}'\mathbf{L} \cdot \sigma^2)$ である。したがって、標準偏差で規準化した統計量

$$Z = \frac{\mathbf{L}'\mathbf{y}}{\sqrt{\mathbf{L}'\mathbf{L} \cdot \sigma^2}}$$

は帰無仮説の下で標準正規分布に従う。

また、 $\mathbf{L}'\mathbf{y}$ とは独立な不偏分散 σ^2 (2.49)を σ^2 に代入した

$$(2.69) \quad t = \frac{\mathbf{L}'\mathbf{y}}{\sqrt{\mathbf{L}'\mathbf{L} \cdot \sigma^2}}$$

は自由度 ν_e のt分布 $t(\nu_e)$ に従う。これからt検定t-testが導かれ、右片側検定の棄却域 R_1 は

$$(2.70) \quad t > t_a(\nu_e)$$

であり、両側検定の棄却域 R_2 は

$$(2.71) \quad |t| > t_{a/2}(\nu_e)$$

となる。

F検定 一般的な次元の帰無仮説についても、BLUEに基づく2次形式統計量からF検定を考えることができる。ここでは次のようにF検定F-testを導く。

まず帰無仮説を

$$(2.72) \quad H_0: E(\mathbf{y}) = \mathbf{X}_0\boldsymbol{\theta}_0$$

のようによく表し

$$(2.73) \quad \text{rank } (\mathbf{X}_0) = p$$

とする。基本モデル(2.9)のランクは

$$(2.74) \quad \text{rank } (\mathbf{X}) = q$$

とする。当然 $p < q$ である。このとき、ランクの差

$$(2.75) \quad s = q - p$$

は基本モデルのうち仮説で零とおかれた部分のランクを表し、仮説の自由度とよばれる。次に基本モデル(2.9)および仮説モデル(2.72)の下でそれぞれ最小二乗法により期待値を推定し、残差平方和 S_e および S_0 を求める。このとき $S_0 > S_e$ である。このとき残差平方和の増分 $S_{H_0} = S_0 - S_e$ が大切である。これに対して次のことが示せる。

$(S_0 - S_e)/\sigma^2$ は H_0 (2.72)の下で自由度 s の χ^2 分布 $\chi^2(s)$ に従う。

[略証] $\mathcal{L}(X), \mathcal{L}(X_0)$ への射影子をそれぞれ Π_{X_0} とおくと

$$(2.76) \quad S_e = \mathbf{y}'(\mathbf{I} - \Pi_X)\mathbf{y},$$

$$(2.77) \quad S_0 = \mathbf{y}'(\mathbf{I} - \Pi_{X_0})\mathbf{y},$$

$$(2.78) \quad S_0 - S_e = \mathbf{y}'(\Pi_X - \Pi_{X_0})\mathbf{y}$$

と表される。ここで、 S_e, S_0 は(4)のようにたがいに独立な規準化変数の二乗和で表されるが、とくに

$$(2.79) \quad S_e = u_1^2 + \cdots + u_{n-q}^2,$$

$$(2.80) \quad S_0 = u_1^2 + \cdots + u_{n-q}^2 + u_{n-q+1}^2 + \cdots + u_{n-q+s}^2$$

$$S_{H_0} = u_{n-q+1}^2 + \cdots + u_{n-q+s}^2$$

と表すことができる(図2.9参照)。これは、また、たがいに共通な変数を含まない S_{H_0} と S_e が独立であることを示している。

(2.79)および(2.80)で u_1, \dots, u_{n-q} はたがいに独立に正規分布 $N(0, \sigma^2)$ に従う。一方、 $u_{n-q+1}, \dots, u_{n-q+s}$ は帰無仮説 H_0 (2.72)が真のときのみ同じように $N(0, \sigma^2)$ に従う。したがって、 S_{H_0}/σ^2 は H_0 の下で $\chi^2(s)$ に従う(終)。

図2.5は最小二乗法の原理であった。この図では平面の中にその部分集合(部分空間)が仮定され、一般的な線形モデルにさらに仮説が加えられていることを表現している。この仮説(部分空間)への距離も考えてそれを元の距離と比べることにより、この仮説を設けることが有意の違いをもたらしたかを見る。

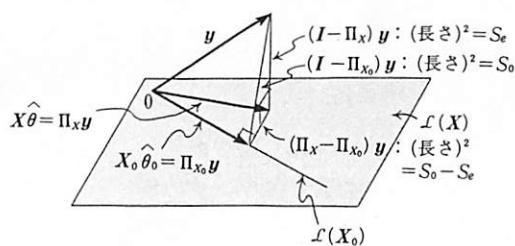


図 2.9 線形検定論の原理：最小二乗法の幾何学的説明
(図 2.5) と比べて何が加わったか理解されたい。

これから H_0 の下で F 分布に従う統計量

$$(2.81) \quad F = \left(\frac{S_0 - S_e}{q-p} \right) / \left(\frac{S_e}{n-q} \right)$$

が定義される。 F は H_0 が真でないときには非心 F 分布 non-central F -distribution に従って、統計的により大きな値をとる。そこで F 分布 $F(q-p, n-q)$ のパーセント点を参照して検定することができる。

例 2.13 〈单回帰モデル： t, F 検定〉 (2.44) における回帰係数 β_1 に関する帰無仮説

$$(2.82) \quad H_0: \beta_1 = 0$$

を検定してみよう。すでに例 2.6 で求めたように

$$\hat{\beta}_1 = S_{xy}/S_{xx}, \quad V(\hat{\beta}_1) = \sigma^2/S_{xx}$$

である。これから t 統計量

$$(2.83) \quad t = \frac{S_{xy}/S_{xx}}{\sqrt{\hat{\sigma}^2/S_{xx}}} = \frac{S_{xy}}{\sqrt{S_{xx}\hat{\sigma}^2}}$$

が得られる。ただし、 $\hat{\sigma}^2$ は (2.57) で与えられる。

帰無仮説 (2.82) は $H_0: E(\mathbf{y}) = \beta_0 \mathbf{j}$ とも表すことができる。この下での残差平方和は、繰返し測定のモデルにほかならないから

$$S_e = \sum y_i^2 - \bar{y}^2/n$$

である。一方、 S_e は (2.56) で与えられるから

$$S_0 - S_e = S_{xy}^2/S_{xx}$$

となる。 $q=2, p=1$ だから $s=q-p=1$ である。以上を (2.81) に代入すると

$$(2.84) \quad F = \left(\frac{S_{xy}^2/S_{xx}}{1} \right) / \left(\frac{S_e}{n-2} \right) \\ = \frac{S_{xy}^2}{S_{xx}\hat{\sigma}^2}$$

が得られる。検定のための 2 つの統計量 (2.83), (2.84) を比べると

$$F = t^2$$

となっていることがわかる。

例 2.14 〈各モデルのデータに対する F の値〉 单回帰モデルのデータ例 2.6' に対して、例 2.9, 2.11 からひき続き計算する。まず、帰無仮説の下での残差平方和、帰無仮説が寄与する増分、基本モデルの下での残差平方和は、それぞれ

$$S_0 = 78,579.3333$$

$$S_0 - S_e = S_{xy}^2/S_{xx} = (377.3467)^2/2.149 = 66,478.6087$$

$$S_e = 12,100.7245$$

となる。この増分の大きさは、 F 統計量

$$F = \frac{66,478.61/1}{12,100.73/13} = 71.4190$$

によって判断される。実際、 $F_{0.05}(1, 13) = 4.667$ であり、増分は有意である。

次に 1 元配置モデル (2.7) における処理の一様性の仮説 (2.67) の検定をしよう。一元配置モデルの残差平方和は

$$S_e = \sum \sum (y_{ij} - \bar{y}_{..})^2 = \sum \sum y_{ij}^2 - \sum y_{..}^2/r$$

で与えられる。これは群内の変動をあらわしている。一方、仮説 (2.67) の下での残差平方和は

$$S_0 = \sum \sum y_{ij}^2 - y_{..}^2/ar$$

となるから

$$S_0 - S_e = \sum y_{..}^2/r - y_{..}^2/ar \\ = r \sum (\bar{y}_{..} - \bar{y}_{..})^2$$

が得られる。これは、群間の変動を表していると考えられる。 $n=ar, q=a, p=1$ より $\nu_e = n-q = ar-a, q-p = a-1$ だから

$$(2.85) \quad F = \frac{r \sum (\bar{y}_{i\cdot} - \bar{y}_{..})^2}{a-1} / \frac{\sum \sum (y_{ij} - \bar{y}_{i\cdot})^2}{a(r-1)}$$

が求める F 統計量である。 (2.85) はその形から群間・群内分散比とよばれている。この考え方は次章の分散分析の考え方と連なってゆく。

これを、データ例 2.2 について計算してみよう。異なる脂肪による処理については

$$\begin{aligned} S_0 &= 64^2 + 72^2 + \cdots + 68^2 - (1770)^2 / 24 \\ &= 3654, \\ S_e &= 64^2 + 72^2 + \cdots + 68^2 - (432^2 + 510^2 + 456^2 + 372^2) / 6 \\ &= 2018, \end{aligned}$$

$$S_0 - S_e = 3654 - 2018 = 1636 \quad (\text{減算})$$

となるから、群間・群内分散比は

$$\begin{aligned} F &= \frac{1636/3}{2018/20} \\ &= 545.3 / 100.9 = 5.40 \end{aligned}$$

となる。これは群間(すなわち処理)が群内に比して F の値で5倍以上かけ離れていることを示している。実際 $F_{0.05}(3, 20) \approx 4.94$ であり、増分は有意である。

2.6.2 F 検定の非心度

t 統計量(2.69)は帰無仮説が成り立たない場合、非心度

$$\gamma = L' \theta / (\sqrt{L' L} \sigma)$$

の非心 t 分布に従う。ただし、非心 t 分布はちょうど非心度に相当する量だけ t 分布をシフトしたものであり、したがって検定の意味は明解である。すなわち、 t 統計量が棄却域にかかる確率(検出力)は自由度 ν_e と非心度 γ のみにより、 γ が 0 から乖離するに従って増大する。

F 統計量(2.85)の場合は、帰無仮説が成り立たないときの分布は非心度

$$(2.86) \quad \gamma = \frac{\theta' X' (\Pi_x - \bar{\Pi}_x) X \theta}{(q-p)\sigma^2}$$

の非心 F 分布であり、 F 検定の検出力も非心度 γ のみによる。

❖ Do not fall in love with a single model! ❖

統計理論の展開上、仮定されるモデルはとかく‘rigid’なものとして扱われがちである。しかしながら、実際のデータに特定の線形モデルが厳密にあてはまり、誤差分布が正確に正規分布従うというようなことはあるはずがない。またそれでよいのであって、統計モデルはあくまでそれを仮定することによって便利に、効率よく推論が実行できるための道具である。

数学理論と異なり、あるデータにはある一つのモデルが唯一無二ということではなく、同じようにあてはまり、同じように有用なモデルがいくつもあり得る。標題はある統計的モデリングに関する専門書からの引用であり、最近のデータ解析の傾向をよく表している。

適切な近似モデルを選ぶ上でもう一つ特記すべきことは、モデルの適合の良さはデータとの距離で測ってはならず、あくまで隠れた真の構造との距離を測るようにすべきことである。うっかりすると手元のデータによくあてはまるモデルを良いモデルと考えがちだが、あくまで背後の母集団をよく記述できるのが良いモデルなのである。

2.6.3 確証的解析と探索的解析

ここまで(正規)線形モデルを仮定し、それを規定するパラメータの推定・検定の方法を一般的に考えてきた。どのような線形モデルが与えられても最小二乗法によりただちに BLUE を構成できるし、興味ある仮説に対して残差平方和の増分を計算することによりただちに F 検定が構成できる。これは、たとえば過去の経験や予備実験から線形モデルが妥当に仮定できる場合に有効であるが、重回帰分析などではしばしば妥当なモデルを探索的に求めことがある。

たとえば、実験の初期の段階では 10~20 もの変数を探り上げ、その中から最終的に少数の重要変数を絞り込むのが一般的である。この場合、回帰係数に関する検定は、あらかじめ設定した特定の仮説を検定するのと異なり、あらゆるモデルの可能性から適切と思われるモデルを選択するという観点から数多く繰り返されることになる。これは同じ線形モデルといつても 1 元配置や 2 元配置が確証的に用いられることが多いのに対し、大きな相違点である。モデル選択の場合に注意すべきことはパラメータ数を増やすと、一見手持ちのデータへのあてはまりはよくなるが、推定量としては不安定で予測式としては役に立た

ないことになりがちなことである。

このことを勘案したモデル選択のため基準として、マローズの C_p 基準、赤池の情報量基準、そしてクロス・バリデーション法などがある。これらの基準はどれも回帰係数の検定で、ほぼ有意水準 15~20%を基準に取捨することに対応する。これに対し確証的な検定では伝統的に有意水準 5%または 1%をとることが多い。これはデータ解析の目的の違いによるものである。

付節 線形代数と統計学^{*)}

統計学は、密度関数、累積分布関数、期待値、分散、積率母関数など、微積分学(解析学)の諸定理や結果を知識として、ふんだんに用いるが、線形代数の知識はそれほど用いない。しかし、統計学は線形代数にとって比較的わかり易い応用領域であって、逆に、統計学への応用によって線形代数がより深く理解されるものである。本章の理解にとって最小限不可欠なもの、あるいはそれを促進する内容のものを、例題を中心にまとめておく。

付1 行列のランクと正規方程式

行列のランク ランクについての最小限の理解を確認しておこう。代数学の多くの問題がそうであるように、これも方程式を解くことから始まる。

i) 連立方程式

$$\begin{cases} 3x - 2y = 7 \\ x + 3y = -5 \end{cases}$$

は、解 $(x, y) = (1, -2)$ をもつ。右辺が他の任意の数 b_1, b_2 で置き換えられても、必ず解があり、クラメールの公式から

$$(*) \quad x = \frac{\begin{vmatrix} b_1 & -2 \\ b_2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & -2 \\ 1 & 3 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} 3 & b_1 \\ 1 & b_2 \end{vmatrix}}{\begin{vmatrix} 3 & -2 \\ 1 & 3 \end{vmatrix}}$$

となる。実際、一次変換(写像)

$$f: \begin{cases} x' = 3x - 2y \\ y' = x + 3y \end{cases}$$

は R^2 (平面全体)を R^2 にうつすから、 x', y' にそれぞれ任意の値 b_1, b_2 をとらせるよう x, y の値を選ぶことができる。

ii) しかし、連立方程式

$$\begin{cases} 3x - 2y = 2 \\ 6x - 4y = 1 \end{cases}$$

は解けない(解が存在しない)。クラメールの公式を使おうとすると、係数の行列式は

$$\begin{vmatrix} 3 & -2 \\ 6 & -4 \end{vmatrix} = 0$$

となって解は見出せない。実際、一次変換

$$g: \begin{cases} x' = 3x - 2y \\ y' = 6x - 4y \end{cases}$$

^{*)} 線形代数学の復習の項なので、習熟している人は読みとばしてよい。

は、 R^2 を R^2 の部分集合(部分空間)

$$l: y' = 2x' \quad (\text{直線})$$

に移すだけだから、連立方程式の右辺がこの l の点でなければ解けない。また、 l の点の場合は解は一意でなくなる。

実際、変換 g は R^2 を1次元空間へ移す。これは、係数の行列

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix}$$

のある性質である。そこで、一般に $m \times n$ 行列 A に対し、一次変換の像

$$\{Ax; x \in R^n\}$$

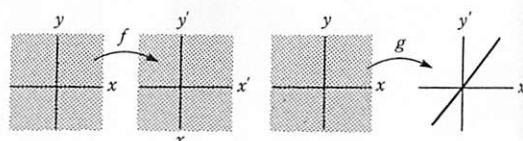
の次元を A のランク(あるいは階数)といい、 $\text{rank}(A)$ で表わす。上例でいえば

$$\text{rank} \begin{pmatrix} 3 & -2 \\ 1 & 3 \end{pmatrix} = 2, \quad \text{rank} \begin{pmatrix} 3 & -2 \\ 6 & -4 \end{pmatrix} = 1$$

である。

i) の変換 f では、2本のベクトル

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \end{pmatrix}$$



が一次独立で、「効果的に働いて」像は2次元になっている。しかし ii)

の変換 g では

$$\begin{pmatrix} 3 \\ 6 \end{pmatrix}, \begin{pmatrix} -2 \\ -4 \end{pmatrix}$$

が、ともに $y=2x$ の同一方向を向いていて、一次従属となっており、無駄ができる。したがって、 g の像は1次元の空間しかなれない(退化している)のである。その意味でランクは、 A の一次独立な列ベクトルの個数ともいいかえられる。

さらに、線型代数学の教える所によれば、これは、 A の小行列式の0でない最大次数ともいいかえてもよい。事実、i) では

$$\begin{vmatrix} 3 & -2 \\ 1 & 3 \end{vmatrix} = 11 \neq 0$$

だから、ランクは2であるが、ii) では

$$\begin{vmatrix} 3 & -2 \\ 6 & -4 \end{vmatrix} = 0, \quad \text{しかし } 3 \neq 0$$

であるから、ランクは1となるのである。

一般に、 $m \times n$ (ただし $m \geq n$)行列 A において、

$$\text{rank}(A)=n$$

のとき、フル・ランク full rank といい、

$$\text{rank}(A) < n$$

のとき、ランク落ちしているといわれる。

とくに正方行列がランク落ちしている場合は、連立方程式の解が存在しても一意ではなく、不定となる。また、やや進んだ知識であるが

$$\text{rank}(A) = \text{rank}(A'A)$$

である。したがって、線形モデル(12.9)の計画行列

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ x_{21} & \cdots & x_{2p} \\ \vdots & & \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \quad (n \times p)$$

のランクが

$$\text{rank}(X) < p$$

ならば、正規方程式の係数行列 $X'X$ ($p \times p$)に対しても

$$\text{rank}(X'X) < p$$

となり、 $X'X$ はランク落ちする。このとき正規方程式の解は不定となり、母数 θ は推定できない。たとえば、標本が小さすぎて

標本の大きさ $n <$ 母数の個数 p

となる場合がこれにあたる。実際、決めるべき対象に比べて情報が少なければ、不定となるのは当然といえる。

付2 最小二乗法と射影子

最小二乗法の理論とその応用は、この射影子の考え方でうまく述べられる。最初は親しみにくいか、線型代数学の具体的理解にも役立ち、その意味や便利さがはっきりしてくるにつれて、最小二乗法にかかわりがあることが理解されてこよう。

射影子 Π の一般的な性質は次の通りである。

- i) $\Pi' = \Pi$ (対称性)
- ii) $\Pi^2 = \Pi$ (べき等性)
- iii) $\text{rank}(\Pi) = \text{tr}(\Pi)$ (ランク(階数)とトレース)
- iii') Π が部分空間 \mathcal{L} への射影子 $\Pi_{\mathcal{L}}$ なら

$$\text{rank}(\Pi_{\mathcal{L}}) = \mathcal{L}$$
の次元,

また $\mathcal{L} = \mathcal{L}(X)$ の場合なら、 $\Pi_{\mathcal{L}} \equiv \Pi_X$ で

$$\text{rank}(\Pi_X) = \text{rank}(X) \quad (\mathcal{L}(X)$$
の次元)

- iv) Π の固有値は $\lambda = 0$ または $\lambda = 1$.

- iv') iv) で、0でない固有値1の個数は $\text{rank}(\Pi)$ に等しい。

これらの i)～iii') を、初等的な解析幾何学の例で見てみよう。

$\langle 2\text{次元} \rangle$ 2次元平面上の点 $P(x, y)$ より直線 $l: y = mx$ への正射影 $Q(x', y')$ は、 p から l への垂線の足であって

$$x' = \frac{x + my}{1 + m^2}, \quad y' = \frac{mx + m^2y}{1 + m^2}$$

で与えられるから、この部分空間 l への射影子は、一次変換

$$\Pi_m = \begin{pmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{pmatrix}$$

となる。これが

$$\begin{aligned}\Pi_m' &= \Pi_m && (\text{対称}) \\ \Pi_m^2 &= \Pi_m && (\text{べき等})\end{aligned}$$

を満すことは見やすい。また

$$|\Pi_m| = 0$$

である。さらにランクは

$$\text{rank}(\Pi_m) = 1$$

である。実際、線形代数学によれば、左辺は Π_m の像 ($\Pi_m x$ の全体) の次元のことであり、その像是直線 l であるから、当然である。

また、 Π_m の対角元の和をとれば、トレースは

$$\text{tr}(\Pi_m) = 1$$

で、 $\text{rank}(\Pi_m)$ に等しい。

数値例として、 $y = mx$ において、 $m = 1$, $m = \sqrt{2}$ とおくと

$$\Pi_1 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad \Pi_{\sqrt{2}} = \begin{pmatrix} 1/3 & \sqrt{2}/3 \\ \sqrt{2}/3 & 2/3 \end{pmatrix}$$

となり、またこれらに直交する直線 $y = -X$, $y = (-1/\sqrt{2})X$ への射影子は、 $m = -1$, $m = -1/\sqrt{2}$ として

$$\Pi_{-1} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}, \quad \Pi_{-1/\sqrt{2}} = \begin{pmatrix} 2/3 & -\sqrt{2}/3 \\ -\sqrt{2}/3 & 1/3 \end{pmatrix}$$

となる。定義から、 Π_1 の像と Π_{-1} の像、 $\Pi_{\sqrt{2}}$ の像と $\Pi_{-1/\sqrt{2}}$ の像は、ベクトルとして直交している。それは、行列の積の計算から

$$\Pi_1 \cdot \Pi_{-1} = O, \quad \Pi_{\sqrt{2}} \cdot \Pi_{-1/\sqrt{2}} = O$$

となり、内積 = 0 となることからも確かめられる。一般に

$$\Pi_m \cdot \Pi_{-1/m} = O$$

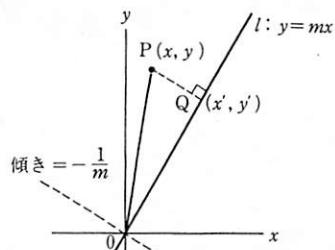
であることを、確認してみるとよい。

ところで、 $\overrightarrow{QP} = \overrightarrow{OP} - \overrightarrow{OQ}$ である。このことを、射影子 Π_1 と点 $P(4, 6)$ について考えてみると、 \overrightarrow{QP} は

$$\begin{pmatrix} 4 \\ 6 \end{pmatrix} - \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

である。また \overrightarrow{QP} は l の垂線だから $y = -X$ の直線(一般には、直交補空間)の中へ、ベクトルとして入っているはずである。事実、そこへの射影子 Π_{-1} を用いると

$$\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



正射影と射影子
直線 l への垂線の足を作れば
正射影が求められる。なお、図
2.5 も参照。

であることが見てとれる。つまり、一般には

$$I - \Pi_m = \Pi_{-1/m} \quad (\text{直交補空間への射影子})$$

が成立している。 $I - \Pi_m$ 自体も射影子であり、対称かつべき等

$$(I - \Pi_m)^2 = I - \Pi_m$$

であることに注意しよう。この式は展開して Π_m のべき等性を用いることすぐに証明できる。

〈3次元〉 3次元空間の例で見ることもできる。

たとえば、 x 軸を含む平面

$$p : z = \sqrt{2}y$$

への射影子は、明らかに x 座標を変えず、 $y-z$ 平面内の射影子だから

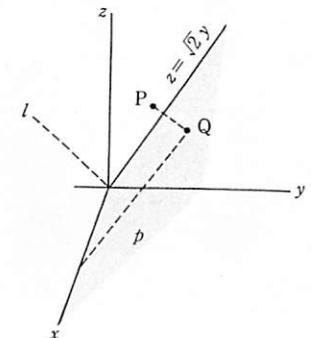
$$\Pi_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & \sqrt{2}/3 \\ 0 & \sqrt{2}/3 & 2/3 \end{pmatrix}$$

となる。これに対しては

$$\text{rank}(\Pi_p) = 2, \quad \text{tr}(\Pi_p) = 2$$

となっていて、やはり両者は相等しい。

また、 p と直交すべき空間、つまり p の直交補空間への射影子は



$$I - \Pi_p = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2/3 & -\sqrt{2}/3 \\ 0 & -\sqrt{2}/3 & 1/3 \end{pmatrix}$$

であって

$$\text{rank}(I - \Pi_p) = 1, \quad \text{tr}(I - \Pi_p) = 1$$

である。実際、いかなる点 (x, y, z) も、この射影子 $I - \Pi_p$ により、原点を通る直線

$$l : x = 0, \quad z = (-1/\sqrt{2})y$$

上へ写されることがわかるが、これは p の垂線 l にほかならない。

〈 n 次元〉 n 次元空間 R^n の点 $P(y_1, y_2, \dots, y_n)$ から、 R^n の直線

$$y_1 = y_2 = \dots = y_n$$

への射影子 Π_l を求める。 P より l へ下した垂線の足

$$(c, c, \dots, c)$$

は、このベクトルが、垂線のベクトル

$$(y_1 - c, y_2 - c, \dots, y_n - c)$$

と直交(内積 = 0)すべきことから

$$Q(\bar{y}, \bar{y}, \dots, \bar{y}) \quad (\bar{y} = \sum y_i/n)$$

と求まる。これが正射影となる。したがって、 l への射影子は

$$\Pi_l = \begin{pmatrix} 1/n & 1/n & \cdots & 1/n \\ 1/n & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 1/n & 1/n & \cdots & 1/n \end{pmatrix}$$

となり、あきらかに

$$\text{rank } (\Pi_l) = 1, \quad \text{tr } (\Pi_l) = 1$$

である。この2つは等しい。

この射影子は統計理論と関係がある。実際、 l の直交補空間への射影子

$$I - \Pi_l = \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1 - \frac{1}{n} \end{pmatrix}$$

は点 P を偏差

$$(y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_n - \bar{y})$$

へ写す。この射影子もランクとトレースが等しく

$$\text{rank } (I - \Pi_l) = n - 1, \quad \text{tr } (I - \Pi_l) = n - 1$$

である(前半は容易に確かめられる)。

偏差が出たので、分散を求めるのにその平方和を

$$\begin{aligned} S &= \sum (y_i - \bar{y})^2 \\ &= ((I - \Pi_l)y)'((I - \Pi_l)y) \\ &= y'((I - \Pi_l)y) \quad (\text{べき等性による}) \end{aligned}$$

のように計算することができる。このように、最小二乗法の原理である‘距離最小’は、射影子の行列演算になっている。このことが、次章以降後々ずっと役立つのである。

付3 固有値問題と残差平方和

行列の固有値問題の例として、例 2.11 で扱った

$$A = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

の固有値、固有ベクトルを求めよう。それは連立方程式

(#)

$$\begin{cases} \frac{2}{3}x - \frac{1}{3}y - \frac{1}{3}z = \lambda x \\ -\frac{1}{3}x + \frac{2}{3}y - \frac{1}{3}z = \lambda y \\ -\frac{1}{3}x - \frac{1}{3}y + \frac{2}{3}z = \lambda z \end{cases}$$

の、 $x=y=z=0$ 以外の解(固有ベクトル)およびそれが存在するための λ の値(固有値)を求ることである。

一般的なやり方にしたがい、まず、固有方程式

$$\begin{vmatrix} \frac{2}{3} - \lambda & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} - \lambda & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} - \lambda \end{vmatrix} = 0$$

を解くと

$$\lambda(\lambda-1)^2=0$$

であるから、固有値は $\lambda=1$ (重解)および $\lambda=0$ となる。

固有値 $\lambda=1$ に対しては、(#)

$$x+y+z=0$$

に帰するから、一般解は、 c_1, c_2 を定数として

$$x=c_1+c_2, \quad y=-c_1+c_2, \quad z=-2c_2$$

すなわち、一次独立な(実は直交する)2つの解

$$(1, -1, 0), \quad (1, 1, -2)$$

の線型結合(の張る2次元平面)となる。ベクトルとしての長さを1に規準化して

$$(1/\sqrt{2}, -1/\sqrt{2}, 0), \quad (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$$

としよう。これが、固有値 $\lambda=1$ に対応する固有ベクトルである。 $\lambda=1$ は2重解なので固有ベクトルは一次独立なものが2個ある。

固有値 $\lambda=0$ に対しては、(#)

$$\begin{cases} 2x - y - z = 0 \\ -x + 2y - z = 0 \end{cases}$$

に帰する。この解は、 c を定数として

$$x=y=z=c$$

あるいは、ベクトル

$$(1, 1, 1)$$

の任意定数倍、すなわち、直線(1次元空間)となる。長さを1に規準化するために

$$(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$$

とすれば、これが固有値 $\lambda=0$ に対応する固有ベクトルである。

付4 多次元正規分布

多くの確率変数を扱うときに、ベクトル、行列表示することが便利であることは、いうまでもない。とくに正規確率変数を扱うときは、便利である以上に多次元確率変数の構造をはっきりさせるうえで、欠かすことができない。実際、本章で扱うこととまとめると次のようになる。

正規線形モデルとは、(2.9)の線形モデル

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

に対して、(2.5)つまり(2.13), (2.14)の諸仮定

$$E(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad V(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$$

が単に $\boldsymbol{\varepsilon}$ の平均、分散の仮定であることに加えて、 $\boldsymbol{\varepsilon}$ の確率分布を n 次元正規分布 $N_n(\mathbf{0}, \sigma^2 \mathbf{I})$

と仮定したモデルである。

ここで多(k)次元あるいは多変量正規分布とは

$$f(\mathbf{x}|\boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

という密度関数をもつ多次元分布である。もし k 個の確率変数の(列)ベクトル \mathbf{X} がこれに従うときには

$$E(\mathbf{X}) = \boldsymbol{\mu}, \quad V(\mathbf{X}) = \Sigma$$

となるので、 $f(\mathbf{x}|\boldsymbol{\mu}, \Sigma)$ は平均ベクトル $\boldsymbol{\mu}$ 、分散・共分散行列 Σ の k 次元正規分布といわれ、 $N_k(\boldsymbol{\mu}, \Sigma)$ と書かれる。

線形モデルでは $\mathbf{X}\boldsymbol{\theta}$ は定数だから、 \mathbf{y} は

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\theta}, \quad V(\mathbf{y}) = \sigma^2 \mathbf{I}$$

となるが、 \mathbf{y} は $\boldsymbol{\varepsilon}$ の線形関数だから正規性は変らず、したがって $N_n(\mathbf{X}\boldsymbol{\theta}, \sigma^2 \mathbf{I})$ に従う。さらに、(\mathbf{X} がフル・ランクの場合)の最小二乗推定量

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

も \mathbf{y} に対する線形演算だから p 次元正規分布に従う。そして

$$\begin{aligned} E(\hat{\boldsymbol{\theta}}) &= E((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{y}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\cdot\mathbf{X}\boldsymbol{\theta} = \boldsymbol{\theta}, \\ V(\hat{\boldsymbol{\theta}}) &= ((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \cdot \sigma^2 \mathbf{I} \cdot ((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X})' \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

である。したがって、最小二乗推定量 $\hat{\boldsymbol{\theta}}$ は、 p 次元正規分布

$$N_p(\boldsymbol{\theta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$$

に従う。これは、(2.42)でも一部求められている。

◆◆◆◆ 第2章 練習問題 ◆◆◆◆

2.1 〈シミュレーション・データの作成〉 $[0, 1]$ 上の一様乱数の列 u_1, u_2, u_3, \dots を発生させ、これから $\varepsilon_i (i=1, 2, \dots)$ を

$$\varepsilon_1 = \sum_{i=1}^{12} u_i - 6, \quad \varepsilon_2 = \sum_{i=13}^{24} u_i - 6, \quad \varepsilon_3 = \sum_{i=25}^{36} u_i - 6, \dots$$

のように定義して、次の確率変数の列 y_i あるいは y_{ij} の値を計算せよ。

i) $y_i = 15 + \varepsilon_i \quad (i=1, 2, \dots, 10)$ (繰返し測定のモデル)

ii) $y_{1j} = 2.0 + \varepsilon_j, \quad y_{2j} = 4.0 + \varepsilon_{5+j}, \quad y_{3j} = 6.0 + \varepsilon_{10+j}, \quad y_{4j} = 8.0 + \varepsilon_{15+j},$
 $y_{5j} = 10.0 + \varepsilon_{20+j} \quad (j=1, 2, \dots, 5)$ (1元配置モデル)

iii) $x_i \equiv i$ として、 $y_i = 10 - 5x_i + 2x_i^2 + \varepsilon_i \quad (i=1, 2, \dots, 10)$ (2次多項式モデル)

なお、i), ii), iii) の各 ε_i は異なるものである。

2.2 〈計画行列〉 前問 2.1 のデータにあてはめるべき繰返しモデル、1元配置モデル、2次多項式モデルの計画行列 X を用意し、 $\mathbf{X}'\mathbf{X}, \mathbf{X}'\mathbf{y}$ を計算せよ。

2.3 〈正規方程式〉 前問 2.1, 2.2 にひき続き、これらのデータに対し各モデルの正規方程式を解き、所期の推定値(2.1において、イタリック体で示されている母数の真の値(に近い値))が得られることを示せ。ただし、逆行列 $(\mathbf{x}'\mathbf{x})^{-1}$ の計算は次のいずれかによるものとする。

- a) 直接、定義による計算を行う方法
- b) 逆行列の数値計算法(掃き出し法など)を学び、それをコンピュータ・プログラミングする方法
- c) コンピュータ・パッケージにある逆行列、最小二乗法などのプログラムを利用する方法

2.4† 〈データの比較〉 以下に示される 4 種類の 2 次元の (x, y) データ・セットがある。 x は制御された変数であり、4 種類のうち 3 種類について共通である。

#(点)	1-3(共通)	1	2	3	4	4
	x	y	y	y	x	y
1	10.0	8.04	9.14	7.46	8.0	6.58
2	8.0	6.95	8.14	6.77	8.0	5.76
3	13.0	7.58	8.74	12.74	8.0	7.71
4	9.0	8.81	8.77	7.11	8.0	8.84
5	11.0	8.33	9.26	7.81	8.0	8.47
6	14.0	9.96	8.10	8.84	8.0	7.04
7	6.0	7.24	6.13	6.08	8.0	5.25
8	4.0	4.26	3.10	5.39	19.0	12.50
9	12.0	10.84	9.13	8.15	8.0	5.56
10	7.0	4.82	7.26	6.42	8.0	7.91
11	5.0	5.68	4.74	5.73	8.0	6.89

(出典: アンスコム)

i) この 4 種類のデータ・セットに対して次の諸量を求めよ。

- a) 観測値の個数,
- b) x の平均 \bar{x} ,
- c) y の平均 \bar{y} ,
- d) 相関係数 r_{xy} ,
- e) 単回帰モデルをあてはめた直線の傾き $\hat{\beta}_1$,
- f) 同、直線,
- g) x の平方和 S_{xx} ,

- h) 偏差積和 S_{xy} , i) 単回帰が寄与する, y の平方和 S_{yy}^2/S_{xx} , j) $H_0: \beta_1=0$ に対応する t 統計量の値 ($\hat{\beta}_1$ の t 値).

ii) 以上から, 何がわかるか.

2.5 <残差の分析> コンクリートの圧縮強度のデータ(表2.4)に対して,

- i) 推定値の標準誤差 standard error of estimates $s.e.=\sqrt{\hat{\sigma}^2}$ を求めよ.

- ii) データの各温度 x_i に対する圧縮強度 y_i の値と, $x=x_i$ に対する y の値の BLUE との差

$$\hat{\epsilon}_i = y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_i), \quad i=1, 2, \dots, 15$$

を計算し, 各 $i=1, 2, \dots, 15$ に対し

$$|\hat{\epsilon}_i| \leq 2 \cdot s.e.$$

が成立しているかどうかを調べよ.

- iii) データ #15 を除いて, 単回帰のモデルをあてはめよ.

- iv) (iii)の解に対し(i), (ii)と同じ計算をせよ. ただし, $i=1, 2, \dots, 14$ とする.

- v) [必要に応じて] データ #1-15 に対して, 2次多項式のモデルをあてはめよ.

2.6[†] <2変量データと回帰分析> 2変量 x, y の間に, 一方(x)が制御された原因, 他方(y)がその結果である場合に, 線形回帰モデルが適切なものとなる. しかし, 変数が制御されていなかったり, 因果関係が明確なものとは考えられなくても, 線形モデルの計算は可能であり, そのような適用は一般に回帰分析とよばれる有効な方法としてよく用いられている.

ある品種のリンゴの樹は, 1本あたりの収量が少ないほどある害虫に損なわれる割合が大きいといわれている. 12本に対する次のデータから, この仮説を検定せよ. また, データを図示せよ.

番号	1本あたり収量 (100個)	虫食いの個数率 (%)
1	8	59
2	6	58
3	11	56
4	22	53
5	14	50
6	17	45
7	18	43
8	24	42
9	19	39
10	23	38
11	26	30
12	40	27

(出典: スネデカー, コクラン)

- 2.7 <同> 溶液中のある物質(y)の沈殿物を得るために試薬(x)が用いられた. データは次の通りである.

x	7.2	4.8	5.2	4.9	5.4	6.4	6.8	8.0	6.0	6.7	7.0	8.0	7.3	4.6	4.2
y	8.4	5.4	6.3	6.8	8.0	11.1	12.3	13.3	8.4	9.5	10.4	12.7	10.3	7.0	5.1

(出典: ボウカ, リーバーマン)

- i) 最小二乗法を用いて直線 $y=\beta_0+\beta_1 x$ を推定せよ.

- ii) $x=6.9$ に対する沈殿物の量の期待値を求めよ.

- iii) $x=6.9$ に対する沈殿物の量の期待値の 95% 信頼限界を求めよ.

- iv) 仮説 $\beta_1=2$ を有意水準 5% で検定せよ.

- 2.8 <同, 2標本> 次のデータは A, B 2 社の社員各 6 名の採用からの年数(x) と給与月額(y , 万円)である.

A社	x	7	12	5	9	11	6
	y	26.7	32.1	25.1	28.5	29.5	25.3

B社	x	4	12	11	8	10	7
	y	25.0	30.7	32.5	28.4	31.8	28.4

(架空例)

A, B 2 社のデータにそれぞれモデル

$$y_{(A)i} = \beta_{0(A)} + \beta_{1(A)}x_{(A)i} + \varepsilon_{(A)i}, \quad \varepsilon_{(A)i} \text{ は } N(0, \sigma_{(A)}^2) \text{ に従う}$$

$$y_{(B)i} = \beta_{0(B)} + \beta_{1(B)}x_{(B)i} + \varepsilon_{(B)i}, \quad \varepsilon_{(B)i} \text{ は } N(0, \sigma_{(B)}^2) \text{ に従う}$$

を想定して, 以下の問い合わせよ.

- i) それぞれの回帰直線と誤差分散の不偏推定値を求めよ.

- ii) 誤差分散が等しいという仮説 $\sigma_{(A)}^2 = \sigma_{(B)}^2$ を有意水準 5% で検定せよ.

- iii) 2つのデータの誤差分散が等しいものとして, 傾きが同じという仮説 $\beta_{1(A)} = \beta_{1(B)}$ を有意水準 5% で検定せよ.

- iv) 2つのデータの誤差分散が等しいものとして, 切片が同じという仮説 $\beta_{0(A)} = \beta_{0(B)}$ を有意水準 5% で検定せよ.

- v) 2つのデータの誤差分散が等しいものとして, 2直線が同一であるという仮説を有意水準 5% で検定せよ.

- 2.9 <続> 2.8 のデータで, A, B 2 社のデータに傾きが共通で切片の異なる直線のモデル

$$y_{(A)i} = \beta_{0(A)} + \beta_{1x(A)i} + \varepsilon_{(A)i}, \quad \varepsilon_{(A)i} \text{ は } N(0, \sigma^2) \text{ に従う}$$

$$y_{(B)i} = \beta_{0(B)} + \beta_{1x(B)i} + \varepsilon_{(B)i}, \quad \varepsilon_{(B)i} \text{ は } N(0, \sigma^2) \text{ に従う}$$

を想定して, 以下の問い合わせよ.

- i) 正規方程式をたて, 2つの直線を求めよ.

- ii) 誤差分散の不偏推定値を求めよ.

iii) 仮説 $\beta_{0(A)} = \beta_{0(B)}$ を有意水準 5%で検定せよ。

2.10 <原点を通る回帰直線> 観測値 y_1, \dots, y_n に対し、原点を通る直線

$$y_i = \beta x_i + \varepsilon_i, \quad i=1, \dots, n$$

を回帰分析のモデルとして考えよう。ただし、 ε_i は $N(0, \sigma^2)$ に従うものとする。このとき

i) β の最小二乗推定量 $\hat{\beta}$ を求めよ。

ii) $\hat{\beta}$ の分布を求めよ。

iii) σ^2 の不偏推定量を求めよ。

iv) 説明変数の特定の値 x における y の平均に対する信頼係数 $1-\alpha$ の信頼限界を求めよ。

第3章 実験データの分析

工学、医学、農学その他多くの分野の実験は

- (1) いくつかの対象や条件の比較を目的としており、
- (2) 同一条件下で実験を繰り返しても結果が同じにならない

という意味で、データに実験誤差が含まれる。という特質をもつ。

この章ではそのような実験データの分析法を説明する。

3.1 2 標本問題

2つの対象の比較 comparison を目的としてデータがとられることはきわめて多い。工業では、製品の品質特性をよくするため、原材料の2つの銘柄や2つの製造方法を比較する実験が行われる。医学でも、薬物の効果や化学物質の毒性を評価するため、次のような実験が行われる。

例 3.1 <対照実験> ある薬物の副作用を調べる目的で、生後3週間のラット20匹をランダムに10匹ずつの2群に分け、一方に薬物を、他方に薬物の入っていない溶媒のみを一定期間投与したのち、体重を測定し、表3.1のデータを得た。

薬物が体重を減少させるといえるか。

表3.1* 2 標本問題データ：ラットの体重(g)

対照群	39.1	40.5	38.0	40.1	46.7	42.3	45.8	36.3	43.2	43.3	(平均 41.53)
処理群	44.7	39.1	37.6	40.0	37.9	34.8	38.4	30.3	40.2	35.2	(平均 37.82)

処理群の10匹には薬物を、対照群の10匹には薬物の入っていない溶媒のみを一定期間投与したのち、体重を測定した。平均の差は $41.53 - 37.82 = 3.71$ であるが、この差が問題となる。

この場合、薬物を与えられた群を処理群 treatment group、そうでない方を対照群 control group とよぶ。また、処理群と対照群を比較するために計画された実験を対照実験 controlled experiment という。こうして、処理群と対照群の2標本を比較する方法が求められる。一般に、2つの条件のもとの観測値の分布を比較する問題を2標本問題 two-sample problem という。この節では、2組のデータ y_{1j} ($j=1, \dots, r_1$), y_{2j} ($j=1, \dots, r_2$) がそれぞれ正規母集団 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ からの独立な観測値であるときの比較について述べよう。

比較の基本となる統計量は、各標本の平均^{*}

$$(3.1) \quad \bar{y}_{1\cdot} = \sum_j y_{1j}/r_1, \quad \bar{y}_{2\cdot} = \sum_j y_{2j}/r_2$$

と平方和

$$(3.2) \quad S_1 = \sum_j (y_{1j} - \bar{y}_{1\cdot})^2, \quad S_2 = \sum_j (y_{2j} - \bar{y}_{2\cdot})^2$$

である。ここで、 $\bar{y}_{1\cdot}$, $\bar{y}_{2\cdot}$ はそれぞれ $N(\mu_1, \sigma_1^2/r_1)$, $N(\mu_2, \sigma_2^2/r_2)$ に従う。また、 S_1/σ_1^2 , S_2/σ_2^2 はそれらと独立にそれぞれ自由度 r_1-1 , r_2-1 の χ^2 分布に従う。母分散 σ_1^2 , σ_2^2 は推定量^{**}

$$(3.3) \quad \hat{\sigma}_1^2 = S_1/(r_1-1), \quad \hat{\sigma}_2^2 = S_2/(r_2-1)$$

で不偏に推定される。

2標本の比較は平均の差 $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ または分散の比 $\hat{\sigma}_1^2/\hat{\sigma}_2^2$ に基づいて行われる。例 3.1 の対照実験では

$$\bar{y}_{1\cdot} = 41.53, \quad S_1 = 100.303, \quad \hat{\sigma}_1^2 = 11.14$$

$$\bar{y}_{2\cdot} = 37.82, \quad S_2 = 132.314, \quad \hat{\sigma}_2^2 = 14.70$$

となっている。具体的方法はいずれ扱う。

3.1.1 母平均の差の検定

2つの母平均が等しいという帰無仮説

$$H_0: \mu_1 = \mu_2$$

^{*}) $y_{i\cdot} = \sum_j y_{ij}$, $\bar{y}_{i\cdot} = \sum_j y_{ij}/r_i$ の “ \cdot ” は、その添字についての和、平均をとったことを意味する(ドット記法)。3.2.2 参照。

^{**}) ハット “ $\hat{\cdot}$ ” は σ_1^2 , σ_2^2 全体に付いている。

を検定する。対立仮説としては通常次のものが考えられる。

両側対立仮説 $H_1: \mu_1 \neq \mu_2$

片側対立仮説 $H_1: \mu_1 > \mu_2$ [または $H'_1: \mu_1 < \mu_2$]

検定の方法は分散の値が既知か未知か、さらに未知のときそれらが等しいかどうかにより 3 つに分かれる。なお、以下ではデータの総数を $n = r_1 + r_2$ とおく。

母分散既知のとき 標本平均の差 $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ は平均 $\mu_1 - \mu_2$, 分散 $(\sigma_1^2/r_1) + (\sigma_2^2/r_2)$ の正規分布に従うので、仮説が正しければ、標準化変数

$$(3.4) \quad z = \frac{\bar{y}_{1\cdot} - \bar{y}_{2\cdot}}{\sqrt{(\sigma_1^2/r_1) + (\sigma_2^2/r_2)}}$$

は標準正規分布に従う。すなわち、差 $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ の評価は (3.4) の分母を基準として行われる。 σ_1^2 , σ_2^2 が既知ならば z が計算できて、有意水準 α の棄却域は

対立仮説が H_1 のとき, $|z| > z_\alpha$

対立仮説が H_2 のとき, $z > z_\alpha$ [H'_1 のとき, $z < -z_\alpha$]

で与えられる。ただし z_α は標準正規分布の上側 α 点である。

母分散未知であるが、等しいと考えられるとき 母分散 σ_1^2 , σ_2^2 が未知でも、 $\sigma_1^2 = \sigma_2^2$ と考えられるときは、その共通の値を σ^2 とおく。いうまでもなく、 $\bar{y}_{1\cdot} - \bar{y}_{2\cdot}$ は平均 $\mu_1 - \mu_2$, 分散 $(1/r_1 + 1/r_2)\sigma^2$ の正規分布に従う。他方、 $(S_1 + S_2)/\sigma^2$ が自由度 $(r_1-1) + (r_2-1) = n-2$ の χ^2 分布に従うので、 σ^2 は併合推定量 pooled variance

$$(3.5) \quad \hat{\sigma}^2 = (S_1 + S_2)/(n-2)$$

で不偏に推定される。したがって、仮説が正しければ

$$(3.6) \quad t = \frac{\bar{y}_{1\cdot} - \bar{y}_{2\cdot}}{\sqrt{\left(\frac{1}{r_1} + \frac{1}{r_2}\right)\hat{\sigma}^2}}$$

の分布は自由度 $n-2$ の t 分布となる。これより、有意水準 α の棄却域は

対立仮説が H_1 のとき, $|t| > t_{\alpha/2}(n-2)$

対立仮説が H_2 のとき, $t > t_\alpha(n-2)$ [H'_1 のとき, $t < -t_\alpha(n-2)$]

で与えられる。ただし $t_\alpha(\nu)$ は自由度 ν の t 分布の上側 α 点である。

例 3.1 の対照実験では、従来の知見から体重は等分散と考えられた。実際、 $\hat{\sigma}_1^2 = 11.14$, $\hat{\sigma}_2^2 = 14.70$ はあまり大きく異っていない(なお、3.1.3 節参照)。

そこでこの方法を用いる。対立仮説は $H_2: \mu_1 > \mu_2$ である。 $\hat{\sigma}^2 = (100.303 + 132.314)/18 = 12.9232$ 、したがって

$$t = (41.53 - 37.82) / \sqrt{(1/10 + 1/10)12.9232} = 2.308$$

となる。 $t_{0.05}(18) = 1.734$ だから、結果は有意水準 5% で有意であり、薬物は体重の減少をもたらすと結論できる。

母分散未知で、等しいとも考えられないとき この場合の取り扱いがもっとも面倒であり、近似的方法が用いられる。(3.4)の右辺の σ_1^2, σ_2^2 を推定量(3.3)で置き換えた

$$(3.7) \quad t' = (\bar{y}_1 - \bar{y}_2) / \sqrt{(\hat{\sigma}_1^2/r_1) + (\hat{\sigma}_2^2/r_2)}$$

の、仮説のもとでの分布は、自由度

$$\frac{((\sigma_1^2/r_1) + (\sigma_2^2/r_2))^2}{\sigma_1^4/r_1^2(r_1-1) + \sigma_2^4/r_2^2(r_2-1)}$$

の t 分布で近似できることが示される。そこで、これらの σ_1^2, σ_2^2 を再び推定量(3.3)で置き換えたものを

$$(3.8) \quad \nu^* = \frac{((\hat{\sigma}_1^2/r_1) + (\hat{\sigma}_2^2/r_2))^2}{\hat{\sigma}_1^4/r_1^2(r_1-1) + \hat{\sigma}_2^4/r_2^2(r_2-1)}$$

とし、 ν を ν^* に最も近い整数として、棄却域を

対立仮説が H_1 のとき、 $|t'| > t_{\alpha/2}(\nu)$

対立仮説が H_2 のとき、 $t' > t_\alpha(\nu)$ [H'_2 のとき、 $t' < -t_\alpha(\nu)$]

とすると、近似的な有意水準 α の検定が得られる。これをウェルチの検定 Welch's test とよぶ。

3.1.2 母平均の差の推定

$\mu_1 - \mu_2$ の不偏推定量は $\bar{y}_1 - \bar{y}_2$ であり、その分散は $(\sigma_1^2/r_1) + (\sigma_2^2/r_2)$ である。区間推定では、検定に対応して 3 通りの場合が生じる。たとえば、信頼係数 $1 - \alpha$ の両側信頼区間の限界はそれぞれ次式で与えられる。

(1) 母分散既知のとき

$$(\bar{y}_1 - \bar{y}_2) \pm z_{\alpha/2} \sqrt{(\sigma_1^2/r_1) + (\sigma_2^2/r_2)}$$

(2) 母分散未知であるが、等しいと考えられるとき

$$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2}(n-2) \sqrt{1/r_1 + 1/r_2} \hat{\sigma}^2$$

※「二重盲検」という工夫※

医薬の効果の検証は、人の健康にかかわるだけに、慎重に行わなければならない。また、効果のない薬を飲みつづけることはそれ自体無駄であり大きな損失である。

しかし、新しい薬が開発されたとき、これが真に効くかどうか科学的に調べることは実は大変である。従来の標準薬と比較するとか、あるいは、9.2 節でもふれるプラセボと比較するということで、対照実験を行って 2 標本検定を行えばよいと、習いたての学生諸君は考えるかもしれない。

しかし、そううまくはいかない。患者は自分に新薬が与えられれば、それだけで‘効く’と心理学的に反応し、快方へ向うかもしれない。一種の「プラセボ」の効果が発生する。かぜなど、自然治癒する病気では、その効果は現実にかなり大きい。その意味では、患者は新薬が与えられたか(処理群)、標準薬ないしはプラセボが与えられたか(対照群)を知らないのが、科学的見地からは適切である。

効果を判定する医者も、新薬に対する期待、あるいは治療業績を求める心理から、新薬に効果ありと判定する傾向をまぬがれ難い。また、学界の‘大先生’の学問的影響も有形・無形にまぬがれないかもしれない。だから、それぞれの患者にどちらが与えられたか、医者も知らないのがよい。

したがって、医者も患者も、2 群への割りあて方を知らず、あるただ一人の第三者(コントローラー)が知っていればよい。これを二重盲検試験 double blind test という。うまい工夫であり、古くはフィッシャーがこれを考案・提案し、今日では、世界はもちろん日本の医薬行政でも、広くとり入れられている。薬効の科学的評価にはたしている役割は大きい。

反面、倫理上の問題もある。患者は医者のくれる薬は可能な限り最良のものと信頼している。信頼違反にならないだろうか。医者は最善の治療を与える職業上の倫理的義務がある。そういう問題は常に残る。

(3) 母分散未知で、等しいとも考えられないとき

$$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/2}(\nu) \sqrt{(\hat{\sigma}_1^2/r_1) + (\hat{\sigma}_2^2/r_2)}$$

ここで用いる自由度 ν は(3.8)の ν^* に最も近い整数である。

3.1.3 分散比の推測

2 つの母分散が等しいという仮説(等分散仮説)

$$H_0: \sigma_1^2 = \sigma_2^2$$

の検定は、分散比そのものに関心がある場合のほか、平均の差の推測において等分散を仮定できるかどうかの判定に用いられることが多い。すでに述べたように S_i/σ_i^2 は自由度 r_i-1 の χ^2 分布に従うので、仮説のもとで

$$(3.9) \quad F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{S_1/(r_1-1)}{S_2/(r_2-1)}$$

の分布は自由度 $\nu_1=r_1-1$, $\nu_2=r_2-1$ の F 分布となる。したがって

対立仮説が $\sigma_1^2 \neq \sigma_2^2$ のとき, $F > F_{\alpha/2}(\nu_1, \nu_2)$ または $F < 1/F_{\alpha/2}(\nu_2, \nu_1)$

対立仮説が $\sigma_1^2 > \sigma_2^2$ のとき, $F > F_\alpha(\nu_1, \nu_2)$

対立仮説が $\sigma_1^2 < \sigma_2^2$ のとき, $F < 1/F_\alpha(\nu_2, \nu_1)$

を棄却域とすることで、有意水準 α の検定が得られる。なお、 $F_{1-\alpha}(\nu_1, \nu_2) = 1/F_\alpha(\nu_2, \nu_1)$ に注意する。

分散比 σ_1^2/σ_2^2 に対する信頼係数 $1-\alpha$ の両側信頼区間は、(3.9)の F から

$$[F/F_{\alpha/2}(\nu_1, \nu_2), \quad F \cdot F_{\alpha/2}(\nu_2, \nu_1)]$$

と表される。

例 3.1 の対照実験では $F = \hat{\sigma}_1^2/\hat{\sigma}_2^2 = 11.14/14.70 = 0.758$ となる。 $F < F_{0.05}(9, 9) = 4.026$, $F > 1/4.026 = 0.248$ だから、有意水準 5% で仮説 $\sigma_1^2 = \sigma_2^2$ は棄却されない。また、信頼係数 95% の両側信頼区間は次のようにある。

$$[0.758/4.026, 0.758 \cdot 4.026] = [0.188, 3.052]$$

なお、この区間が 1 を含むことからも、等分散の仮説は棄却されないことがわかる。

3.2 1元配置分散分析

3.2.1 1元配置とそのモデル

2つの母集団平均 μ_1, μ_2 を比較する方法は 3.1 節で説明した。3つ以上の母集団平均 $\mu_1, \mu_2, \dots, \mu_a$ ($a \geq 3$) の比較 (k 標本問題 k -sample problem) には以下の分散分析 analysis of variance, 略して ANOVA が基本的な役割を果たす。

例 3.2 〈化学反応の反応条件〉 化学工場で、製品中のある成分の含有量を多くするため、反応温度を

3.2 1元配置分散分析

表 3.2† 1元配置実験データ：製品 1 kg 中の成分量(g)

反応温度	A_1 (50°C)	A_2 (55°C)	A_3 (60°C)	A_4 (65°C)	
デ タ	77.4	78.3	79.2	78.9	
	78.2	78.2	79.3	78.8	
	78.1	78.4	79.1	78.1	
	77.8	77.3	78.2	78.1	
計	389.4	391.3	395.1	392.8	総計 1568.6
平均	77.88	78.26	79.02	78.56	総平均 78.43

反応温度を 4 段階に変え、各反応温度で 5 回ずつ製品をつくり、製品 1 kg 中の成分量(g)を調べた。反応温度という因子の 4 水準が横方向に配置されている。縦方向は同一条件における各 5 回の繰り返しであって、縦・横の 2 方向に因子が配置されている。後述の例 3.3 の表 3.5(2 元配置)とは異なる。

このデータでは、各水準内のデータの変動に比べて水準平均には大きな差が観察される。この差が統計的に有意か否かの結論は分散分析表 3.3 で与えられる。

$$A_1: 50^\circ\text{C}, \quad A_2: 55^\circ\text{C}, \quad A_3: 60^\circ\text{C}, \quad A_4: 65^\circ\text{C}$$

の 4 段階に変え、つくられた製品 1 kg 中の含有量を調べた。各反応温度で 5 回ずつ実験を行った結果を表 3.2 に示す。反応温度によって含有量は変化するかを考えよう。

実験結果に影響をおよぼすと考えて実験で取り上げるもの因子 factor および、因子に対して与える条件を水準 level とよぶ。この実験は反応温度という単一の因子をもつ 1 因子実験であり、この因子は $A_1: 50^\circ\text{C}, A_2: 55^\circ\text{C}, A_3: 60^\circ\text{C}, A_4: 65^\circ\text{C}$ の 4 つの水準をもつ。2 因子以上の場合は各因子の水準の組合せが実験されることになるが、このような具体的な実験条件のことを処理

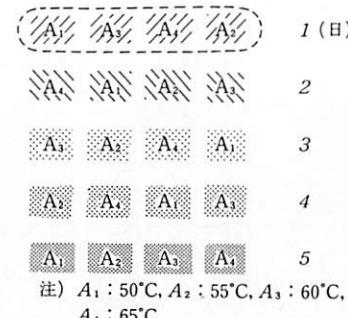


図 3.1 ブロックの考え方(点線内)
各 1 日の中では条件は同一であるから、その中の A_1 と A_4 の比較は、系統的な誤差をもたらす‘日’という因子に左右されないと考えられる。各ブロック(日)の中での順序は、乱数表で決めればよい。

treatment とよぶ。なお、各水準での繰り返しは同じ回数でなくてもよい。

処理の比較を目的とする実験では、因子として取り上げていないさまざまな要因による系統的な誤差が比較に偏りを生じないように配慮する必要がある。そのため、これらの要因の影響が確率的に公平になるよう、第11章の乱数表などを用いて実験順序をランダムにすることが重要である。この例では1日4回の実験が可能で、5日間にわたって行われた20回の実験全部の順序がランダマイズ(無作為化、確率化)された。このような実験法を完全無作為化法 completely randomized designといい、この方法による1因子実験を1元配置 one-way layout という。

さて、実験計画法が生まれた農事試験では、試験区の地力等の変動が大きいため、全体をいくつかのブロックに分け、その内部ができるだけ均一になるように管理したうえで、品種や施肥法といった比較したい処理をその中にランダムに割り当てた。このように、実験の場をブロック block とよばれる内部の比較的均一ないくつかの部分に分け、各ブロックで処理の一揃いを実験する方法を乱塊法 randomized block design という(ランダムを‘乱’、ブロックを‘塊’とした)。

この例でも、天候などに起因する成分含有量の日間変動が存在するときは、日をブロックとし、毎日4つの水準をランダムな順序で実験する乱塊法を用いる方が、日間の差異が実験誤差から分離されるので、反応温度の比較精度はよくなるであろう。同様に、多数の装置やオペレータがあるときは、そのうちのいくつかを選んでブロックとすることが考えられる。

乱塊法で得られた実験データの分析法はこの節で述べるものとは異なる。

いま、因子を A_i 、その a 個の水準を A_{i1}, \dots, A_{ia} 、各水準での繰り返し数を r_1, \dots, r_a とする。 A_i 水準の j 番目のデータを y_{ij} と表し、データの構造として次のモデルを想定する。

$$(3.10) \quad y_{ij} = \mu_i + \varepsilon_{ij}, \quad i=1, 2, \dots, a; j=1, \dots, r_i$$

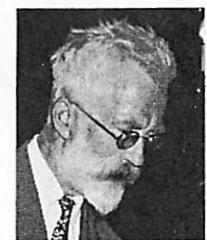
ここに母数 μ_i は第 i 水準に固有な平均であり、それに加わる実験誤差 ε_{ij} はすべて互いに独立に $N(0, \sigma^2)$ に従うものとする。ただし、この共通な分散 σ^2 は未知である。

以上から、例3.2は正規母集団 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2), \dots, N(\mu_a, \sigma^2)$ の平均の比較の問題となる。前節の2標本問題では、分散に関する事前情報によって異なった3つの検定法を用いたが、1元配置の場合は上のように等分散を仮定する。この仮定は後述のバートレットの検定(3.22)でチェックできる。

※大統計学者の生れ育った土壤※

物理学や化学の実験では、実験条件をできるだけ精密に管理することによって実験誤差を小さくすることに注意が払われる。しかし、農事試験では、酸・アルカリ度(pH)のような土壤の性質などの条件を完全に一定にすることは不可能であり、実験結果のばらつきは避けられない。また、作物が生育するまでに一定の時間がかかることから、実験回数をいくらでも大きくするわけにはいかない。しかも、実験農場の規模では小さな収量の差も、国土全体ではきわめて大きな差になる。こうして、限られた回数の実験から得られた誤差を含んだデータから、収量の差をできるだけ正確に偏りなく見いだすことが農事試験の大きな課題であり、そこに実験計画法の生まれる文字通り‘土壤’があったといえよう。

「近代統計学の父」といわれる R. A. フィッシャーは、1912年にケンブリッジ大学を卒業したあと、パブリック・スクールなどで働いたのち、蓄積された実験結果を研究するため1919年にロザムステッド農事試験場の職に就いた。そして、そこで実験計画法や分散分析法をはじめ、現在の統計的推測理論の基本的な枠組みを構築する数々の重要な研究を精力的に進めたのである。(本項98頁へ続く)



R. A. フィッシャー

いま、データ総数を $n = \sum r_i$ とし、繰り返し数 r_i の重みでの μ_i の荷重平均

$$\mu = \sum r_i \mu_i / n$$

を一般平均 grand mean とよぶ。全水準に共通な平均である。表3.2でいえば、どのような反応温度でもほぼこれくらいの成分含有量が得られるという値になる。各水準の平均からこれを引いたものが、その正味の効果となる。こうして、

$$\alpha_i = \mu_i - \mu, \quad i=1, 2, \dots, a$$

を A_i 水準の効果 effect とよぶ。なお、 $\sum r_i \alpha_i = 0$ が成立することに注意する。この μ と α_i を用いると、(3.10)は

$$(3.11) \quad y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i=1, 2, \dots, a; j=1, \dots, r_i$$

と表される。これは(共通の効果)+(第 i 水準の効果)+(それ以外の誤差)という形になっている。

例3.2の反応条件のデータにおいては、 μ は総平均で推定されると考えられ

るから、その推定値は $\hat{\mu}=78.43$ (g)となる。また、 $50^\circ, 55^\circ, 60^\circ, 65^\circ\text{C}$ に対しては、おのおの

$$\hat{\mu}_1=77.88, \quad \hat{\mu}_2=78.26, \quad \hat{\mu}_3=79.02, \quad \hat{\mu}_4=78.56 \quad (\text{g})$$

であり、正味の効果はこれらの $\hat{\mu}$ からの偏差

$$\hat{a}_1=-0.55, \quad \hat{a}_2=-0.17, \quad \hat{a}_3=0.59, \quad \hat{a}_4=0.13 \quad (\text{g})$$

となる。

3.2.2 分散分析

データに関する想定(3.10)あるいは(3.11)を1元配置 one-way layout のモデルとよぶ。分散分析は、例3.2で求められているように、因子 A のすべての水準の平均が等しいという帰無仮説

$$H_0: \mu_1=\mu_2=\cdots=\mu_a$$

あるいは、これと同等な帰無仮説

$$H_0: \alpha_1=\alpha_2=\cdots=\alpha_a=0$$

を検定する方法である。

繰り返し数が一定な1元配置モデルに関する推測は第2章の一般線形モデルで説明した。繰り返し数が一定でない場合の結果は以下のようになる。

なお、ドット記法

$$y_{i\cdot}=\sum_j y_{ij}, \quad y_{..}=\sum_i \sum_j y_{ij} \quad (\text{表3.2の計, 総計欄})$$

$$\bar{y}_{i\cdot}=y_{i\cdot}/r_i, \quad \bar{y}_{..}=y_{..}/n \quad (\text{表3.2の平均, 総平均欄})$$

を用いてある。

データに1元配置モデル(3.10)をあてはめたときの残差平方和は

$$(3.12) \quad S_e = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2 \\ = \sum_i \sum_j y_{ij}^2 - \sum_i y_{i\cdot}^2 / r_i$$

であり、 S_e/σ^2 は自由度 $v_e=n-a$ の χ^2 分布に従う。また、仮説 $H_0: \mu_1=\mu_2=\cdots=\mu_a$ のもとでの、つまり、モデル $y_{ij}=\mu+\varepsilon_{ij}$ をあてはめたときの残差平方和は、 μ を総平均 $\bar{y}_{..}$ で推定して

$$(3.13) \quad S_T = \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = \sum_i \sum_j y_{ij}^2 - y_{..}^2 / n$$

となる。こうして、仮説 H_0 を設けたことによる残差平方和の増加分は

$$(3.14) \quad S_A = S_T - S_e \\ = \sum_i y_{i\cdot}^2 / r_i - y_{..}^2 / n = \sum_i r_i (\bar{y}_{i\cdot} - \bar{y}_{..})^2$$

となる。

S_A は S_e と独立で、仮説が正しければ自由度 $v_A=a-1$ の χ^2 分布に従う。したがって、仮説のもとで

$$(3.15) \quad F = \frac{S_A/v_A}{S_e/v_e}$$

が自由度 v_A, v_e の F 分布 $F(v_A, v_e)$ に従うので、このことを用いて仮説が検定できる。これが分散分析検定 analysis of variance test, ANOVA test である。

平方和の分解 (3.15)は次のように考えると理解しやすい。 $\mu_1, \mu_2, \dots, \mu_a$ が等しいかどうかは、それらの推定値 $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_a$ の変動の大きさから判断される。例3.2でいえば、それは

$$S_A = 5\{(77.88-78.43)^2 + (78.26-78.43)^2 + (79.02-78.43)^2 \\ + (78.56-78.43)^2\} \\ = 5\{(-0.55)^2 + (-0.17)^2 + 0.59^2 + 0.13^2\} = 3.482$$

で表される。この値を判断する基準となるのが、各水準内の変動(誤差変動)

$$S_e = (77.4-77.88)^2 + (78.2-77.88)^2 + \cdots + (77.9-77.88)^2 + \cdots \\ + \cdots + (78.9-78.56)^2 + (78.8-78.56)^2 + \cdots + (78.9-78.56)^2 \\ = 3.620$$

である。こうして、平均が等しいという仮説 H_0 の検定が(3.15)の分散の比に基づいて行われることになる。「分散分析」という語の由来である。

さて、データ全体の変動の大きさは、総平方和*とよばれる(3.13)の S_T によって表現されると考えられる。総平方和を上記の「水準変更による変動」 S_A と「同一条件下の変動」 S_e に分解しよう。そのため、 S_T の各成分を

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{i\cdot} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i\cdot})$$

と表し、両辺を2乗して加えると、交差項の和は

$$\sum_i \sum_j (\bar{y}_{i\cdot} - \bar{y}_{..})(y_{ij} - \bar{y}_{i\cdot}) = \sum_i ((\bar{y}_{i\cdot} - \bar{y}_{..})) \sum_j (y_{ij} - \bar{y}_{i\cdot}) = 0$$

となるから

* 総平方和を総変動 total variation ともよぶ。後述の誤差平方和、級間平方和も同様である。

$$(3.16) \quad \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = \sum_i r_i (\bar{y}_{i..} - \bar{y}_{..})^2 + \sum_i \sum_j (y_{ij} - \bar{y}_{i..})^2$$

が成立する。

いうまでもなく、(3.16)の右辺第1項が S_A 、第2項が S_e である。 S_A は水準の変更にともなうデータの変動の大きさを表し、級間平方和(因子 A を明示するときは A 間平方和)とよばれる。 S_e は同一実験条件でのデータの変動の大きさを表し、誤差平方和とよばれる。(3.16)の分解

$$(3.16') \quad S_T = S_e + S_A$$

を平方和の分解とよぶ。これは(3.14)と同じである。

平方和の計算 分散分析検定のための平方和 S_A, S_e の計算は、(3.16)の定義通りに行うよりも、修正項 correction term とよばれる

$$CT = \bar{y}_{..}^2/n$$

を最初に求め

$$S_T = \sum_i \sum_j y_{ij}^2 - CT, \quad S_A = \sum_i y_{i..}^2 / r_i - CT, \quad S_e = S_T - S_A$$

の順に求める方がずっと簡単である。

例 3.2 の反応条件のデータでは、表 3.2 から $\bar{y}_{..} = 1568.6$ 、したがって

$$CT = (1568.6)^2 / 20 = 123025.298$$

$$S_T = 77.4^2 + 78.2^2 + \cdots + 78.9^2 - 123025.298 = 123032.40 - 123025.298 \\ = 7.102$$

$$S_A = (389.4^2 + 391.3^2 + 395.1^2 + 392.8^2) / 5 - 123025.298 \\ = 123028.78 - 123025.298 = 3.482$$

$$S_e = 7.102 - 3.482 = 3.620$$

となる。

この計算はパソコンの倍精度で行ったものであり、通常の桁数の電卓ではできない。この例からもわかるように、この計算法では、 $123032.40 - 123025.298$ のように、大きさのあまり違わない 2 数の差として平方和が求められるため、桁落ちが生じることが多い。たとえば、有効数字 6 桁で上記の計算を行えば、最初の 1 桁も正確かどうかわからない。

これを防ぐには、事前にデータに近い一定の数——仮平均とよぶ——を全データから引き去っておくのがよい。いまの場合、たとえば仮平均を 77 とし、77.4 を 0.4

に、78.2 を 1.2 に、……というようにデータを変換しておくのである。このようにしても平方和は変化しない。また、水準平均の推定では、こうして計算された平均に仮平均をたして元にもどしてやればよい。

分散分析表 平方和 S_A と S_e は、それらを構成している変数の個数が多くなれば、大きな値をとりやすくなることはいうまでもない。比較に際してはそれを考慮しなければならない。それが(3.15)である。

各平方和の自由度は、平方和に含まれる独立な成分の個数に一致する。まず、 S_A の自由度 ν_A は $a-1$ である。実際、 S_A は $\bar{y}_{1..} - \bar{y}_{..}, \dots, \bar{y}_{a..} - \bar{y}_{..}$ という a 個の成分をもつが、 $\sum_i r_i (\bar{y}_{i..} - \bar{y}_{..}) \equiv 0$ であるから、それらのうちの $(a-1)$ 個の値を決めれば最後の 1 値は決まってしまう。次に、 $S_e = \sum_i \sum_j (y_{ij} - \bar{y}_{i..})^2$ は $n = \sum_i r_i$ 個の成分からなる平方和であるが、 $i=1, \dots, a$ に対して $\sum_j (y_{ij} - \bar{y}_{i..}) \equiv 0$ であるから、 $y_{11} - \bar{y}_{1..}, \dots, y_{ir_i} - \bar{y}_{i..}$ のうち独立な成分の個数は $r_i - 1$ 、したがって S_e の自由度 ν_e はそれらの和である $n-a$ になる。

周知のように S_T の自由度は $n-1$ だから、 $\nu_T = n-1, \nu_A = a-1, \nu_e = n-a$ となり、平方和の分解(3.16')に対応して、自由度についても

$$\nu_T = \nu_A + \nu_e$$

が成り立つことがわかる。ここで

$$(3.17) \quad V_A = S_A / \nu_A, \quad V_e = S_e / \nu_e$$

とおき*、それぞれ因子 A および誤差の平均平方 mean square とよぶ。これを用いると、分散分析検定(3.15)の検定統計量は

$$(3.18) \quad F = V_A / V_e$$

と表される。

(3.11) から $\bar{y}_{i..} = \mu + \alpha_i + \bar{\varepsilon}_{i..}$, $\bar{y}_{..} = \mu + \bar{\varepsilon}_{..}$ となるので

$$y_{ij} - \bar{y}_{i..} = \varepsilon_{ij} - \bar{\varepsilon}_{i..}, \quad \bar{y}_{i..} - \bar{y}_{..} = \alpha_i + \bar{\varepsilon}_{i..} - \bar{\varepsilon}_{..}$$

したがって

$$S_e = \sum_i \sum_j (\varepsilon_{ij} - \bar{\varepsilon}_{i..})^2, \quad S_A = \sum_i r_i (\alpha_i + \bar{\varepsilon}_{i..} - \bar{\varepsilon}_{..})^2$$

が得られる。この表現から、水準効果 α_i の大きさに関する情報は S_A に含まれ

*）標本分散は S^2, s^2 を用いるのが通常であるが、本章では分散分析法での慣用に従う。

$E(S_A) = (a-1)\sigma^2 + \sum_i r_i \alpha_i^2$
となること、および、 S_e は誤差の大きさに関する情報を担い
 $E(S_e) = (n-a)\sigma^2$
が成立することが示される。

のことから

$$E(V_A) = \sigma^2 + \sum_i r_i \alpha_i^2 / (a-1), \quad E(V_e) = \sigma^2$$

が成り立つ。仮説 $H_0 : \mu_1 = \mu_2 = \dots = \mu_a$ は $\sum_i r_i \alpha_i^2 = 0$ と同等なので、 H_0 のもとでは $E(V_A) = \sigma^2$ となり、 F は 1 に近い値をとることが期待される。一方、 H_0 が正しくなければ F は大きな値をとり易い。

検定は、 H_0 のもとで F が自由度 ν_A, ν_e の F 分布に従うことを利用して行われる。これで分散分析の手続きが終わる。

例 3.2 の反応条件のデータでは、 $\nu_A = 4-1=3, \nu_e = 20-4=16$ 、したがって

$$V_A = 3.482/4 = 1.1607, \quad V_e = 3.620/16 = 0.2263$$

となる。これから

$$F = 1.1607 / 0.2263 = 5.13$$

が得られる。 $F_{0.05}(3, 16) = 3.24$ だから結果は有意水準 5% で有意であり、成分含有量は反応温度によって変化するという結論となる。

分散分析の結果は、表 3.3 の形の分散分析表 analysis of variance table, ANOVA table にまとめるのが通例である。ここで、 F 値が有意水準 5% で有意のときは*印を 1 つ、1% で有意のときは*印を 2 つつけるのが慣習になっている。

なお、仮説 $H_0 : \mu_1 = \mu_2 = \dots = \mu_a$ を $H_{01} : \mu_1 = \mu_2, H_{02} : \mu_2 = \mu_3, \dots, H_{0,a-1}$:

表 3.3[†] 反応条件データ(例 3.2, 表 3.2)の分散分析表

要 因	平 方 和	自 由 度	平均 平 方	F 値
反応温度(A)	$S_A = 3.482$	$\nu_A = 3$	$V_A = 1.1607$	$F = 5.13^*$
誤 差	$S_e = 3.620$	$\nu_e = 16$	$V_e = 0.2263$	
計	$S_T = 7.102$	$\nu_T = 19$		

$F = 5.13$ が $F_{0.05}(3, 16) = 3.24$ より大きいので、結果は 5% 有意であり、 H_0 は棄却され、成分量が反応温度によって変化するという結論となる。

$\mu_{a-1} = \mu_a$ という $a-1$ 個の仮説に分解して、それを 3.1 節の t 検定を用いて検定すれば H_0 が検定できるのではないかと考える人がいる。しかし、それぞれの検定の有意水準を α とすると、仮説 H_0 が正しいにもかかわらず $H_{01}, H_{02}, \dots, H_{0,a-1}$ の少なくともひとつを棄却する確率は α より大きくなってしまうので、そのような方法は適切でない。

3.2.3 処理効果の推定

因子の効果が有意となった場合、通常は水準平均に関する推定が行われる。第 i 水準の平均 μ_i は、その水準のデータの平均で推定される。すなわち

$$\hat{\mu}_i = \bar{y}_i.$$

このとき、 $E(\bar{y}_{i..}) = \mu_i, V(\bar{y}_{i..}) = \sigma^2/r_i$ だから、信頼係数 $1-\alpha$ の信頼限界は

$$(3.19) \quad \bar{y}_i \pm t_{\alpha/2}(\nu_e) \sqrt{V_e/r_i}$$

で与えられる。また、2つの水準平均の差 $\mu_i - \mu_j$ の点推定値は

$$\hat{\mu}_i - \hat{\mu}_j = \bar{y}_i - \bar{y}_j.$$

であり、これに対する信頼係数 $1-\alpha$ の信頼限界は次のようになる。

$$(3.20) \quad \bar{y}_i - \bar{y}_j \pm t_{\alpha/2}(\nu_e) \sqrt{\left(\frac{1}{r_i} + \frac{1}{r_j}\right)V_e}$$

例 3.2 の反応条件のデータにおいて、平均が最大となった A_3 水準(60°C)の平均に対する 95% 信頼限界は、 $\nu_e = 20-4=16, t_{0.05}(16)=2.120$ より

$$79.02 \pm 2.120 \sqrt{0.2263/5} = 79.02 \pm 0.45 = 78.57, 79.47$$

となる。各水準平均に 95% 信頼区間を付したものを見図 3.2 に示した。

また、 $\mu_3 - \mu_4$ に対する 95% 信頼区間の限界は

$$79.02 - 78.56 \pm 2.120 \sqrt{0.2263 \times 2/5} = 0.46 \pm 0.64 = -0.18, 1.10$$

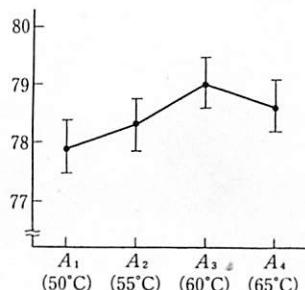


図 3.2 例 3.2 の反応条件データからの
水準平均の推定値と信頼区間
本文に述べた検定により 60°C と 65°C の平均に有意差がないことが示される。

となる。この区間が0を含むので、 μ_3 と μ_4 に、つまり 60°C と 65°C の平均に有意な差がないと結論できる。

3つ以上の分散の一様性の検定 分散分析で仮定された分散の一様性を検定する方法はいくつか提案されているが、代表的なバートレットの検定を紹介する。それは、各水準の不偏分散を

$$V_i = \sum_j (y_{ij} - \bar{y}_{i..})^2 / (r_i - 1), \quad i=1, \dots, a$$

それらを併合したものを

$$V_e = \sum_i V_i / (n-a) = \sum_i \sum_j (y_{ij} - \bar{y}_{i..})^2 / (n-a)$$

とするとき ($n = \sum_i r_i$)、 \log を自然対数として

$$(3.21) \quad B = (n-a) \log V_e - \sum_i (r_i - 1) \log V_i$$

が、等分散仮説 $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_a^2$ のもとで近似的に自由度 $a-1$ の χ^2 分布に従うことを用いるものである。このとき、 χ^2 近似をよくするために

$$(3.22) \quad B' = \frac{B}{1 + \frac{1}{3(a-1)} \left\{ \sum_i \frac{1}{r_i-1} - \frac{1}{n-a} \right\}}$$

と補正する。

例3.2の反応条件のデータで等分散仮説を検定してみよう。 $r_i-1=4$, $n-a=16$ であり

$$V_1=0.0970, \quad V_2=0.4130, \quad V_3=0.2170, \quad V_4=0.1780, \quad V_e=0.2263$$

より $B=2.103$ 。これと、 $1 + \{4/(5-1) - 1/(20-4)\} / \{3(4-1)\} = 1.104$ から、 $B'=2.103/1.104=1.905$ となる。 $\chi^2_{0.05}(3)=7.81$ であるから等分散仮説は棄却されない。

3.3 交互作用と要因実験

例3.2の実験では反応温度だけを因子としてとりあげた。実際には、原料の銘柄、触媒量など、収量に影響すると考えられるものは数多く存在する。2つ以上の因子をとりあげる実験を一般に多因子実験とよぶ。多因子実験の場合は、各因子の単独の効果(主効果 main effect という)ばかりでなく、それらの組合せによる効果にも注意を払う必要がある。

2因子 A, B の水準組合せでの応答が表3.4(a)のような場合、 B_1 と B_2 の差が A の水準によらずすべて20と一定なので、これを図示した図3.3(a)では

表3.4(a) 交互作用がない場合の応答

	B_1	B_2
A_1	30	50
A_2	60	80
A_3	70	90

A の各水準での B_1 と B_2 の差が一定である。 B の各水準についても同じ。

表3.4(b) 交互作用がある場合の応答

	B_1	B_2
A_1	30	40
A_2	60	90
A_3	70	50

A の各水準での B_1 と B_2 の差が一定でない。 B の各水準についても同じ。

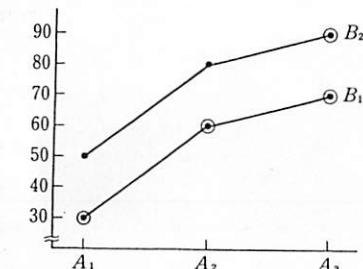


図3.3(a) 交互作用がない場合の応答

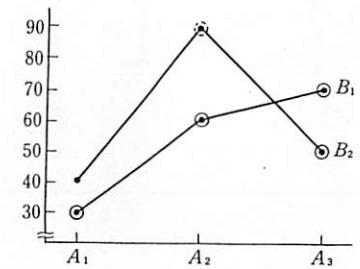


図3.3(b) 交互作用がある場合の応答

B_1 と B_2 の差は、 A の水準 A_1, A_2, A_3 に左右されず、応答を結ぶ線分が平行になる。単一因子実験では、 A_1B_1, A_2B_1, A_3B_1 をまず比較して、因子 A の水準を A_3 と決定した後、 A_3B_1, A_3B_2 を比較して最適な水準組合せ A_3B_2 に到達する。

B_1 と B_2 の差は、 A の水準 A_1, A_2, A_3 によって異なり、応答を結ぶ線分が平行にならない。

(a)の場合に行なったような単一因子実験を行うと A_3B_1 に達し、最適な水準組合せ A_2B_2 に到達できない。

応答を結ぶ線分が平行になっている。この場合、 A_1B_1 から A を A_2 に変更したときの応答の増加分30と、 B を B_2 に変更したときの増加分20の和が、これらを同時に変更したときの増加分50に等しく、水準変更にともなう応答の変化に加法性が成立する。このような場合、2因子 A, B に交互作用 interaction が存在しないという。

一方、表3.4(b)を図示した図3.3(b)の場合、応答を結ぶ線分は平行でない。 B_1 と B_2 の差は A の水準に依存している。そして、 A_1B_1 から A を A_2 に変更したときの増加分30と、 B を B_2 に変更したときの増加分10の和40は、それらを同時に変更したときの増加分60に一致しない。このような場合、2因子 A, B 間に交互作用が存在するという。交互作用の大きさの定義は、この

後の「2元配置分散分析」で具体的に与える。

多因子の最適な水準組合せを求めるとき、しばしば次のような実験が行われる。最初に、 A 以外の因子 B, C, \dots の水準を適当に固定したうえで、 A の各水準 A_1, A_2, \dots を比較する実験を行う。その結果に基づいて A の最適水準を決定し、以下の実験ではつねにその水準を採用する。同様にして B, C, \dots と順に最適水準を決定し、それらを組み合わせて最適な水準組合せとする。このような実験法を单一因子 one-factor-at-a-time 実験という。これとは異なり、因子のすべての水準組合せで実験する方法を要因実験 factorial experiment という。

単一因子実験では、交互作用が存在するとき、最適な水準組合せに到達できるとはかぎらないし、なによりも交互作用の有無をチェックすることができない。いま、 A, B に交互作用がある図 3.3(b)のような場合に、単一因子実験を用いて応答が最大になる水準組合せを求めるものとする。最初に B を B_1 に固定したものとすれば、 A_1B_1, A_2B_1, A_3B_1 の比較から A_3 が最適水準として選択され、次に A_3B_1 と A_3B_2 の比較から A_3B_1 が最適な水準組合せと結論される(実際に実験される組合せを◎で示した)。本当の最大値を与えるのは A_2B_2 だから、これは正しくない。

単一因子実験でも、交互作用が存在しない図 3.3(a)のような場合は最大値に到達する。しかし、たとえば B_1 と B_2 の差に関して単一の推定値しか得られないのに、すべての組合せを実験する要因実験では、 A_1, A_2, A_3 のそれぞれで 3 つの独立な推定値が得られる。実際の実験では観測値に誤差が含まれるので、実験回数が 4 から 6 に増加するだけで 3 倍の推定値が得られる要因実験の方が効率的である。

実験回数があまり多くならなければ要因実験を用いることが薦められる。

3.4 2元配置分散分析

多因子要因実験を完全無作為化法で行うとき、それを多元配置とよび、2 因子の場合を 2 元配置 two-way layout, 3 因子の場合を 3 元配置等という。多元配置データから交互作用を含めた各因子の効果を推測する方法を、最も簡単な 2 元配置の場合について説明しよう。

例 3.3 〈金属の圧延条件〉 金属の圧延工場で、圧延製品の引張り強さ

表 3.5* 繰り返しのある 2 元配置データ：圧延製品の引張り強さ (kg/mm²)

速度 率 率	B_1	B_2	B_3	B_4
A_1	76.0	76.2	75.3	70.4
	73.2	79.8	79.9	72.4
A_2	77.1	84.5	86.5	76.9
	75.0	80.7	83.1	78.1
A_3	75.7	81.3	80.7	77.1
	74.7	76.9	83.4	82.0

圧延率($A : 3$ 水準), 圧延速度($B : 4$ 水準)の 2 因子を選び、繰り返し $r=2$ として、24 回の実験をランダムな順序で行った。

表 3.2 の 1 元配置の場合と異なり、この表では因子が縦横 2 方向に配置され、その交点に各水準組合せでの繰り返し 2 のデータが記入されている。

(kg/mm²)を高めるため、因子として圧延率($A : 3$ 水準), 圧延速度($B : 4$ 水準)の 2 つを選んで実験した。繰り返し数を $r=2$ とし、全部で 24 回の実験はランダムな順序で行った。その結果を表 3.5 に示す。圧延率や圧延速度は引張り強さに影響するか、さらには最適な水準組合せはなにかを考えよう。

3.4.1 2元配置のモデル

因子 A, B の水準組合せ A_iB_j での k 番目の観測値を y_{ijk} とする。なお、 A, B の水準数をそれぞれ a, b とし、繰り返し数は一定で r とする。繰り返し数が一定でないデータの分析は容易ではない。例 3.3 の圧延条件のデータでは $a=3, b=4, r=2$ である。この y_{ijk} の構造として次のモデルを想定する。

$$(3.23) \quad y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad i=1, 2, \dots, a; \quad j=1, \dots, b; \quad k=1, \dots, r$$

ただし μ_{ij} は水準組合せ A_iB_j の平均であり、実験誤差 ε_{ijk} は互いに独立に $N(0, \sigma^2)$ に従うものとする。データの総数を $n=abr$ とおく。

ひとまず、 A の第 i 水準 A_i, B の第 j 水準 B_j 、および全体の「効果」を

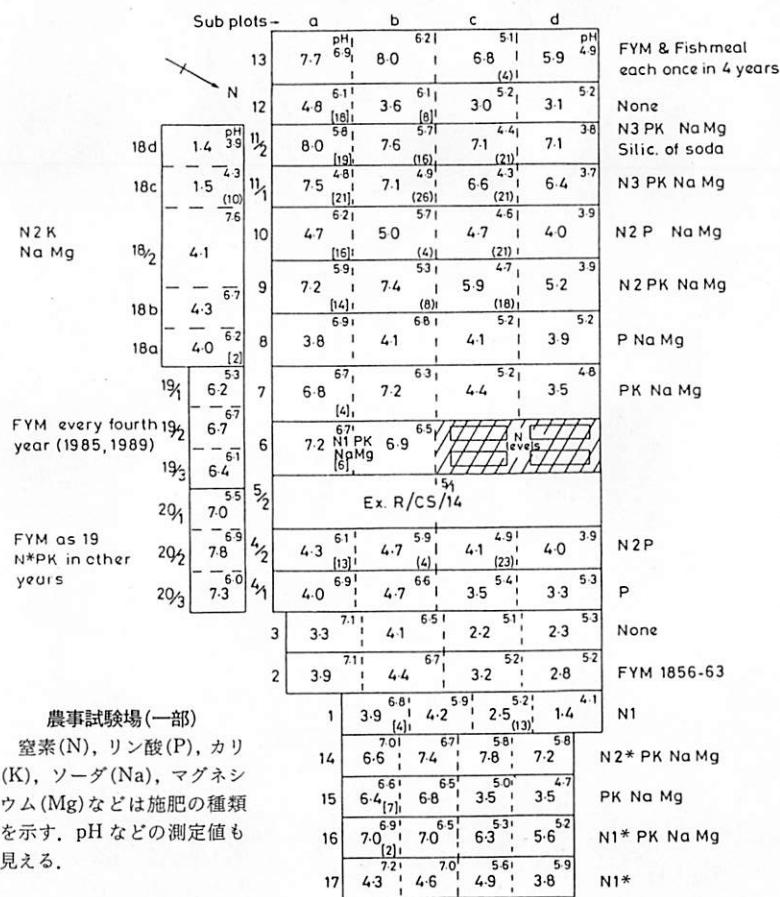
$$\bar{\mu}_i = \sum_j \mu_{ij}/b, \quad \bar{\mu}_{.j} = \sum_i \mu_{ij}/a, \quad \mu = \sum_i \sum_j \mu_{ij}/ab$$

とおき、 μ を一般平均とよぶ。そして A_i, B_j の「効果」から μ を引いた正味

$$\alpha_i = \bar{\mu}_i - \mu, \quad i=1, \dots, a; \quad \beta_j = \bar{\mu}_{.j} - \mu, \quad j=1, \dots, b$$

❀さすがは統計学を生んだ国❀

ロザムステッド農事試験場は、ロンドンの北の郊外ハーベンデンにある世界最古の農業研究所で、その規模もまた世界有数である。John Bennet Lawes (1814-1900) は、青年時代に培った化学に対する関心をやがて農業に向け、1843年父祖の地にこの農事試験場を設立した。そして、その生涯を通じて土壤の化学的組成や施肥が作物の生育に与える影響を調べるための実験を組織的、継続的に行った。ここから近代統計学も育っていった。（日本ではまだ江戸時代であった。）図は1856年末のロザムステッドの試験状況である。



をそれぞれ因子 A, B の主効果 main effect とよぶ.

さらに、 μ_{ij} のうち主効果の重ね合わせと一般平均とで表せない部分^{*)}

$$(\alpha\beta)_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) = \mu_{ij} - \bar{\mu}_{i\cdot} - \bar{\mu}_{\cdot j} + \mu \quad i=1, 2, \dots, a; j=1, \dots, b$$

を因子 A, B の交互作用 interaction とよぶ。もし $(\alpha\beta)_{ij} \equiv 0$ なら

$$\mu_{ij} \equiv \mu + \alpha_i + \beta_j$$

で、表 3.4(a)の加法性が成り立ち、交互作用はないことになる。

これらを用いると(3.23)は

$$(3.24) \quad y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

のように、(一般平均) + (因子 A の効果) + (因子 B の効果) + (因子 AB の交互作用) + (誤差)という形に表される。

なお、主効果や交互作用には

$$(3.25) \quad \sum \alpha_i = 0, \quad \sum \beta_j = 0;$$

$$\sum_i (\alpha\beta)_{ij} = 0, \quad j=1, \dots, b; \quad \sum_j (\alpha\beta)_{ij} = 0, \quad i=1, \dots, a$$

が成立することに注意しよう。

3.4.2 平方和の分解

表3.6は、表3.5の圧延条件のデータについて、水準組合せ A_iB_j の合計 y_{ij} 、水準 A_i の合計 $y_{i..}$ および平均 $\bar{y}_{i..}$ 、水準 B_j の合計 $y_{.j..}$ および平均 $\bar{y}_{.j..}$ 、さらに総計 $y_{...}$ および総平均 $\bar{y}_{...}$ を示している。

この表から、(3.25)の μ , α_i , β_j , $(\alpha\beta)_{ij}$ の値の見当(推定値)が得られる。ま
ず、一般平均については

表 3.6† 2 元配置データの平方和計算のための要約表：例 3.3 の圧延条件データ

	B_1	B_2	B_3	B_4	$y_{i..}$	$\bar{y}_{i..}$
A_1	149.2	156.0	155.2	142.8	603.2	75.40
A_2	152.1	165.2	169.6	155.0	641.9	80.24
A_3	150.4	158.2	164.1	159.1	631.8	78.98
$y_{.j.}$	451.7	479.4	488.9	456.9	1876.9($y_{...}$)	
$\bar{y}_{.j.}$	75.28	79.90	81.48	76.15		78.20($\bar{y}_{...}$)

表 3.5 のデータを要約したるものである。

*) ここで ' $a\beta$ ' は積でなく一つの文字と考えられたい。必要なら γ を用いて γ_{β} とすればよい。

$$\hat{\mu} = \bar{y} \dots = 78.20$$

となり、また $\mu_{1\cdot}$ は $\hat{\mu}_{1\cdot} = \bar{y}_{1\cdot\cdot} = 75.40$, $\mu_{\cdot 1}$ は $\hat{\mu}_{\cdot 1} = \bar{y}_{\cdot 1\cdot} = 75.28$ のように推定される。これから主効果は、 A_1, B_1 に対しては

$$\hat{\alpha}_1 = \bar{y}_{1\cdot\cdot} - \bar{y} \dots = -2.80, \quad \hat{\beta}_1 = \bar{y}_{\cdot 1\cdot} - \bar{y} \dots = -2.92$$

などのように推定される。交互作用は、 A_1B_1 についての $(\alpha\beta)_{11}$ ならば、 $\hat{\mu}_{11} = \bar{y}_{11\cdot} = 149.2/2 = 74.6$ だから

$$\hat{(\alpha\beta)}_{11} = \bar{y}_{11\cdot} - \bar{y}_{1\cdot\cdot} - \bar{y}_{\cdot 1\cdot} + \bar{y} \dots = 2.12$$

と推定される。その他の母数も同様に推定できる。これらの推定値には(3.25)の関係が成り立つ。

われわれの目的は処理の比較だから、一般平均には関心がない。そこで、全データから一般平均を差し引いておくと

$$\begin{aligned} (3.26) \quad y_{ijk} - \bar{y} \dots &= (\bar{y}_{ij\cdot} - \bar{y} \dots) + (y_{ijk} - \bar{y}_{ij\cdot}) \\ &= (\bar{y}_{ij\cdot} - \bar{y} \dots) + (\bar{y}_{\cdot j\cdot} - \bar{y} \dots) \\ &\quad + (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y} \dots) + (y_{ijk} - \bar{y}_{ij\cdot}) \end{aligned}$$

という恒等式が成立する。最右辺の第1, 2, 3項はそれぞれ既に説明した $\hat{\alpha}_i, \hat{\beta}_j, \hat{(\alpha\beta)}_{ij}$ である。最後の項は処理 A_iB_j の平均からの偏差であり、同一条件での実験の繰り返しによる誤差(実験誤差)に対応する項である。したがって、これが(3.24)のモデルに対応するデータでの表現になる。

データ全体の変動の大きさを表す総平方和 S_T は、(3.26)の第2辺の表現に対応して

$$S_T = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y} \dots)^2 = r \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y} \dots)^2 + \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij\cdot})^2$$

と分解される。交差項の和が0になることは1元配置の場合と同様に示せる。右辺第1項は因子 A, B の水準変更にともなうデータの変動の大きさを表し、 S_{AB} と表される。第2項は同一実験条件でのデータの変動の大きさを表す誤差平方和 S_e である。

この S_{AB} を、(3.26)の最右辺の表現に対応して

$$\begin{aligned} S_{AB} &= r \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y} \dots)^2 \\ &= br \sum_i (\bar{y}_{i\cdot\cdot} - \bar{y} \dots)^2 + ar \sum_j (\bar{y}_{\cdot j\cdot} - \bar{y} \dots)^2 + r \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y} \dots)^2 \end{aligned}$$

と分解する。この右辺の各項を順にそれぞれ A 間平方和 S_A , B 間平方和 S_B , 交互作用平方和 $S_{A\times B}$ とよぶ。こうして、(3.26)から、2元配置における最終

的な平方和の分解

$$(3.27) \quad S_T = S_A + S_B + S_{A\times B} + S_e$$

が得られる。

平方和の計算 各平方和の計算は、修正項

$$CT = y \dots^2/n, \quad (\text{ただし}, n = abr)$$

を最初に求め

$$\begin{aligned} S_T &= \sum_i \sum_j \sum_k y_{ijk}^2 - CT, & S_{AB} &= \sum_i \sum_j y_{ij\cdot}^2 / r - CT, \\ S_e &= S_T - S_{AB}, \\ S_A &= \sum_i y_{i\cdot\cdot}^2 / br - CT, & S_B &= \sum_j y_{\cdot j\cdot}^2 / ar - CT, \\ S_{A\times B} &= S_{AB} - S_A - S_B \end{aligned}$$

の順に求めるのがよい。

例3.3の圧延条件のデータでは、 $y_{ij\cdot}, y_{i\cdot\cdot}, y_{\cdot j\cdot}, y \dots$ は表3.6に与えられている。したがって

$$CT = 1876.9^2/24 = 146781.400$$

$$S_T = 76.0^2 + 73.2^2 + \dots + 82.0^2 - 146781.400 = 372.910$$

$$S_{AB} = (149.2^2 + 156.0^2 + \dots + 159.1^2)/2 - 146781.400 = 308.175$$

$$S_e = 372.910 - 308.175 = 64.735$$

$$S_A = (603.2^2 + 641.9^2 + 631.8^2)/8 - 146781.400 = 100.736$$

$$S_B = (451.7^2 + 479.4^2 + 488.9^2 + 456.9^2)/6 - 146781.400 = 158.278$$

$$S_{A\times B} = 308.175 - 100.736 - 158.278 = 49.161$$

となる。この場合も、事前に全データから仮平均(たとえば80)を引いておいて桁落ちを防ぐのがよい。

3.4.3 分散分析表

1元配置の場合に述べたように、平方和 $S_A, S_B, S_{A\times B}$ の大きさを S_e を基準として判断するには、それぞれの自由度を考慮しなければならない。1元配置のときの考え方から、各平方和の自由度は

$$\begin{aligned} (3.28) \quad \nu_T &= n-1, & \nu_A &= a-1, & \nu_B &= b-1, \\ \nu_{A\times B} &= (a-1)(b-1), & \nu_e &= ab(r-1) \end{aligned}$$

となる ($\nu_{A \times B} = \nu_A \times \nu_B$ に注意)。平方和の分解に対応して次式が成立している。

$$(3.29) \quad \nu_T = \nu_A + \nu_B + \nu_{A \times B} + \nu_e$$

各平方和を自由度で割って、平均平方を求める。

$$(3.30) \quad V_A = S_A / \nu_A, \quad V_B = S_B / \nu_B,$$

$$V_{A \times B} = S_{A \times B} / \nu_{A \times B}, \quad V_e = S_e / \nu_e$$

容易に示されるように、各平方和の期待値について

$$(3.31) \quad E(S_A) = (a-1)\sigma^2 + br \sum_i \alpha_i^2$$

$$E(S_B) = (b-1)\sigma^2 + ar \sum_j \beta_j^2$$

$$E(S_{A \times B}) = (a-1)(b-1)\sigma^2 + r \sum_i \sum_j (\alpha\beta)_{ij}^2$$

$$E(S_e) = ab(r-1)\sigma^2$$

が成立する(練習問題3.4)。したがって、各平均平方の期待値は

$$(3.32) \quad E(V_A) = \sigma^2 + br \sum_i \alpha_i^2 / (a-1),$$

$$E(V_B) = \sigma^2 + ar \sum_j \beta_j^2 / (b-1),$$

$$E(V_{A \times B}) = \sigma^2 + r \sum_i \sum_j (\alpha\beta)_{ij}^2 / (a-1)(b-1)$$

$$E(V_e) = \sigma^2$$

と表される。

(3.32)から、各仮説^{*}の検定に用いられるF統計量は

$$H_0 : (\alpha\beta)_{ij} \equiv 0 \text{ の検定は, } F_{A \times B} = V_{A \times B} / V_e$$

$$H_0 : \alpha_i \equiv 0 \text{ の検定は, } F_A = V_A / V_e$$

$$H_0 : \beta_j \equiv 0 \text{ の検定は, } F_B = V_B / V_e$$

となる。

第1の仮説は交互作用がないというものであり、1元配置には見られなかつたものである。交互作用が存在しなければ、一方の因子の水準変更は他の因子の水準によらず主効果の差だけの変化をもたらす。また、各水準組合せの平均 μ_{ij} は主効果の重ね合せで表されるので、最適な水準組合せを求めるときも、それぞれの因子の最適水準を組み合わせればよい。

しかし、交互作用が存在する場合は上に述べたことは成り立たない。とくに、 μ_{ij} の挙動は主効果の重ね合せで記述できないので、最適な水準組合せを上述の方法で求めることはできない。こうして、交互作用が有意なときは主効果に

^{*}) ' $\alpha_i \equiv 0$ ' は $\alpha_1 = \dots = \alpha_a = 0$ を意味する。 $\beta_j \equiv 0, (\alpha\beta)_{ij} \equiv 0$ についても同じ。

表3.7[†] 圧延条件データ(例3.3, 表3.5)の分散分析表

要 因	平 方 和	自 由 度	平均 平 方	F 値
圧延率 (A)	$S_A = 100.736$	$\nu_A = 2$	$V_A = 50.368$	$F_A = 9.34^{**}$
圧延速度 (B)	$S_B = 158.278$	$\nu_B = 3$	$V_B = 52.759$	$F_B = 9.78^{**}$
交互作用 (A × B)	$S_{A \times B} = 49.161$	$\nu_{A \times B} = 6$	$V_{A \times B} = 8.194$	$F_{A \times B} = 1.52$
誤 差	$S_e = 64.735$	$\nu_e = 12$	$V_e = 5.395$	
計	$S_T = 372.910$	$\nu_T = 23$		

$F_{A \times B} = 1.52 < F_{0.05}(6, 12)$ だから交互作用(A × B)は有意でない。 $F_A = 9.34 > F_{0.01}(2, 12) = 6.93, F_B = 9.78 > F_{0.01}(3, 12) = 5.95$ だから圧延率(A), 圧延速度(B)はともに1%有意である。

このように、2元配置分散分析表では、2つの因子の他に交互作用が変動要因として現れる。

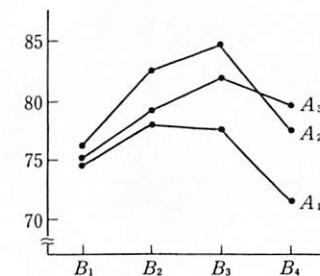


図3.4[†] 例3.3の圧延条件のデータの各水準組合せでの平均
平均を結ぶ線分に極端な非平行性がみられないことは交互作用がないことを示唆する。
また、最適な水準組合せは A_2B_3 のように見える。

あまり積極的な意味がなくなる。したがって、交互作用の仮説を最初に検定しなければならない。

例3.3の圧延条件のデータに対する結果は表3.7の分散分析表にまとめられる。交互作用は有意でなく、A: 圧延率、B: 圧延速度がともに1%有意である。水準組合せの平均のグラフを図3.4に与えた。

3.4.4 処理効果の推定

各水準組合せにおける平均の推定法は、交互作用が有意かどうかで異なる。

交互作用が有意なとき、水準組合せ A_iB_j での平均 μ_{ij} は A_iB_j でのデータの平均で推定するほかはない。すなわち

$$(3.33) \quad \hat{\mu}_{ij} = \bar{y}_{ij}.$$

とする。これに対する信頼係数 $1-\alpha$ の信頼限界は次のとおりである。

$$(3.34) \quad \bar{y}_{ij} \pm t_{\alpha/2}(\nu_e) \sqrt{V_e/r}$$

交互作用が有意でないとき、さらに主効果のどちらか、たとえば B が有意でなければ、モデルは

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} = \mu + \alpha_i + \varepsilon_{ijk}$$

となる。したがって、 A に関する1元配置で、各水準に br 回の繰り返しがある場合に帰着する。また、主効果が両方とも有意な場合、 μ_{ij} は

$$\mu_{ij} = \mu + \alpha_i + \beta_j = (\mu + \alpha_i) + (\mu + \beta_j) - \mu$$

と表される。この最右辺の各項をそれぞれ $\bar{y}_{..}$, $\bar{y}_{.j}$, $\bar{y}_{...}$ で推定し、

$$(3.35) \quad \hat{\mu}_{ij} = \bar{y}_{..} + \bar{y}_{.j} - \bar{y}_{...}$$

としてやればよい。この場合でも推定量(3.33)を用いることができるが、その分散 σ^2/r は(3.35)の分散 $\sigma^2(a+b-1)/abr$ より大きい。(3.35)を用いたときの信頼係数 $1-\alpha$ の信頼限界は

$$(3.36) \quad \hat{\mu}_{ij} \pm t_{\alpha/2}(v_e) \sqrt{V_e(a+b-1)/(abr)}$$

である。処理平均の差の推定については練習問題を参照のこと。

例 3.3 の圧延条件のデータでは交互作用が有意でないので、最適な水準組合せは A, B の最適水準を組み合わせた A_2B_3 である。 μ_{23} の推定には(3.35)を用いる。表 3.6 から

$$\begin{aligned} \hat{\mu}_{23} &= \bar{y}_{2..} + \bar{y}_{.3} - \bar{y}_{...} \\ &= 80.24 + 81.48 - 78.20 = 83.52 \end{aligned}$$

これに対する 95% 信頼限界は

$$\begin{aligned} &83.52 \pm 2.179 \sqrt{5.395(3+4-1)/24} \\ &= 83.52 \pm 2.53 = 80.99, 86.05 \end{aligned}$$

となる。

3.4.5 繰り返しのない2元配置

後で述べるように、繰り返しのない2元配置では、交互作用と実験誤差の大きさを分離して把握することができない。したがってそれは、従来の知見などによって交互作用が存在しないことがわかっているときに用いられる。

例 3.4 <熱処理歪み> 金属製品の熱処理にともなう変形(熱処理歪み)を小さくする目的で、材料の種類(A : 3 水準)と、熱処理の温度と時間を適当に組み合わせた熱処理条件(B : 4 水準)を因子とする実験を行った。技術上の理由から交互作用が存在しないと考えられたので、繰り返しのない2元配置を用い、12回の実験をランダムな順序で行って変形量(mm)を測定し、表 3.8 の

3.4 2元配置分散分析

表 3.8* 繰り返しのない2元配置データ：熱処理歪み(mm)

材料	条件	B_1	B_2	B_3	B_4	$y_{..}$	$\bar{y}_{..}$
A_1		3.4	3.9	3.7	3.5	14.5	3.625
A_2		3.3	3.6	3.8	3.8	14.5	3.625
A_3		3.0	3.3	3.5	3.4	13.2	3.300
$y_{.j}$		9.7	10.8	11.0	10.7	42.2(y..)	
$\bar{y}_{.j}$		3.233	3.60	3.667	3.567		3.517(\bar{y}..)

材料(A : 3 水準)と熱処理条件(B : 4 水準)の 2 因子を選び、全部で 12 回の実験をランダムな順序で行った。技術上の理由から交互作用は存在しないと考えられたので、繰り返しのない2元配置を用いた。

データを得た。材料の種類や熱処理条件は変形量に影響するかを考えよう。水準組合せ A_iB_j でのデータ y_{ij} の構造に関し、次のモデルを想定する。

$$(3.37) \quad \begin{aligned} y_{ij} &= \mu_{ij} + \varepsilon_{ij} \\ &= \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad i=1, 2, \dots, a; j=1, \dots, b \end{aligned}$$

ここで、主効果 α_i, β_j は条件(3.25)を満たし、実験誤差 ε_{ij} は互いに独立に $N(0, \sigma^2)$ に従うものとする。(3.37)に対応するデータでの表現は

$$y_{ij} - \bar{y}_{..} = (\bar{y}_i - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_i - \bar{y}_{.j} + \bar{y}_{..})$$

となる。これから

$$(3.38) \quad S_T = S_A + S_B + S_e$$

ただし、

$$(3.39) \quad \begin{aligned} S_T &= \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2, \quad S_A = b \sum_i (\bar{y}_i - \bar{y}_{..})^2, \quad S_B = a \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 \\ S_e &= \sum_i \sum_j (y_{ij} - \bar{y}_i - \bar{y}_{.j} + \bar{y}_{..})^2 \end{aligned}$$

という平方和の分解が得られる。

平方和の計算は、まず修正項

$$CT = \bar{y}_{..}^2 / (ab) = 42.2^2 / 12 = 148.4033$$

を求め、

$$\begin{aligned} S_T &= \sum_i \sum_j y_{ij}^2 - CT = 149.14 - 148.4033 = 0.7367 \\ S_A &= \sum_i y_i^2 / b - CT \\ &= (14.5^2 + 14.5^2 + 13.2^2) / 4 - 148.4033 = 0.2817 \\ S_B &= \sum_j y_{.j}^2 / a - CT \\ &= (9.7^2 + 10.8^2 + 11.0^2 + 10.7^2) / 3 - 148.4033 = 0.3367 \end{aligned}$$

$$S_e = S_T - S_A - S_B \\ = 0.7367 - 0.2817 - 0.3367 = 0.1183$$

の順に求めるのがよい。

各平方和の自由度は

$$\nu_T = ab - 1, \quad \nu_A = a - 1, \quad \nu_B = b - 1, \quad \nu_e = (a-1)(b-1)$$

となる。各平方和を自由度で割って、平均平方を求める

$$(3.40) \quad V_A \equiv S_A/\nu_A, \quad V_B \equiv S_B/\nu_B, \quad V_C \equiv S_C/\nu_C$$

各仮説の検定に用いられる F 統計量は

$H_0: \alpha_i \equiv 0$ の検定は、 $F_A \equiv V_A/V_{eA}$

$H_0: \beta_i \equiv 0$ の検定は、 $F_B = V_B/V_a$

となる。こうして、表3.9の分散分析表を得る。結果はA:材料、B:熱処理条件とともに有意水準5%で有意である。

分散分析後の処理効果に関する推定については、繰り返しのある2元配置に比べて目新しい点はないので、ここでもう一度説明することはしない。以下では繰り返しのない2元配置について2つのことを述べておこう。

第1に、交互作用が存在するとき、繰り返しのない2元配置のデータの構造モデルは、(3.37)ではなく

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij}$$

となり、交互作用と誤差が同じ添字をもつことになる。その結果、(3.39)の S_e には両方の大きさに関する情報が混じり合って含まれることになり、両者を分離して把握することはできなくなってしまう。したがって、分散分析検定は正しい結果を与えない可能性がある。

表 3.9† 热处理歪みデータ(例 3.4, 表 3.8)の分散分析表

要因	平方和	自由度	平均平方	F 值
材料(A)	$S_A = 0.2817$	$\nu_A = 2$	$V_A = 0.1408$	$F_A = 7.14^*$
熱処理条件(B)	$S_B = 0.3367$	$\nu_B = 3$	$V_B = 0.1122$	$F_B = 5.69^*$
誤差	$S_e = 0.1183$	$\nu_e = 6$	$V_e = 0.0197$	
計	$S_T = 0.7367$	$\nu_T = 11$		

$F_A = 7.14 > F_{0.05}(2, 6) = 5.14$, $F_B = 5.69 > F_{0.05}(3, 6) = 4.76$ だから、材料(A), 熱処理条件(B)はともに 5% 有意である

繰り返しのある 2 元配置の分散分析表に比べ、交互作用の項がないので、それだけ単純な形になる

第2に、3.2節で述べた1因子乱塊法実験は、もともとの因子をA、プロックを因子Bとし、繰り返しのない2元配置のデータとして分析することができる。ただし、ブロック効果が有意になった場合、処理効果に関する推定は繰り返しのある2元配置から類推されるものとは違ってくる。これについては実験計画法のテキストにまかせよう。

3.4.6 2水準直交表

多因子実験では水準組合せのすべてを実験する要因実験の方が単一因子実験より有利である。しかし、よくばった結果因子数が多くなると水準組合せの数、したがって要因実験に必要な実験回数はきわめて多くなる。たとえば、因子が10個あれば、すべてが2水準でも水準組合せの数は $2^{10}=1024$ となり、少なくとも1024回の実験が必要になる。これは通常実施できない。

企業における新製品開発のための研究などでは、製品の特性に影響する要因が技術的に絞り込めないことが多いので、実験で取り上げる因子数が多くなることは珍しくない。そこで、時間や費用の点から、もっと少ない回数で必要な情報が得られるような実験計画はできないか、という要請が出てくる。それに答えるのが直交表による実験である。

表 3.10 は 2 水準直交表の例である。中央部に数字 1 と 2 が並んでおり、それらは各列に同じ回数(4 回)現れる。さらに、7 つの列のどの 2 列を取り上げても $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$ という 4 つの組合せが同じ回数(2 回)現れる。

表 3.10[†] 2 水準直交表

実験番号	列番	1	2	3	4	5	6	7	水準組合せ	データ
1		1	1	1	1	1	1	1	$A_1B_1C_1D_1E_1$	20
2		1	1	1	2	2	2	2	$A_1B_1C_1D_2E_2$	21
3		1	2	2	1	1	2	2	$A_1B_2C_2D_1E_1$	25
4		1	2	2	2	2	1	1	$A_1B_2C_2D_2E_2$	24
5		2	1	2	1	2	1	2	$A_2B_1C_2D_1E_2$	29
6		2	1	2	2	1	2	1	$A_2B_1C_2D_2E_1$	27
7		2	2	1	1	2	2	1	$A_2B_2C_1D_1E_2$	23
8		2	2	1	2	1	1	2	$A_2B_2C_1D_2E_1$	25
対応させる因子		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>				

実験の回数が飛躍的に節約される。

いま、実験に取り上げる5個の因子 A, B, C, D, E がすべて2水準であるとする。このとき、表の7列から適当に5列を選んで、それらに5因子を対応させ、表の数字が各因子の水準を示すものとして8回の実験の水準組合せを決定する。たとえば、対応を表の最下行のようにすれば、表の右に示したように、実験番号1の水準組合せは $A_1B_1C_1D_1E_1, \dots$ 、実験番号8の水準組合せは $A_2B_2C_1D_2E_1$ となる。

各因子の第2水準と第1水準の差は、第2水準と第1水準の実験データの平均の差で推定される。たとえば因子Aの場合、実験番号1~4が第1水準、5~8が第2水準だから、上のデータから

$$(29+27+23+25)/4 - (20+21+25+24)/4 = 3.5$$

が差の推定値となる。直交表の上記の性質から、実験番号1~4と5~8には因子 B, C, D, E の2つの水準が同じ回数現れるので、それらとAとの交互作用がなければ、この推定値は他の因子の効果の影響を受けない。

この場合、すべての水準組合せを実験すると $2^5=32$ 回の実験が必要になる。

◆◆◆◆ 第3章 練習問題 ◆◆◆◆

3.1 <2標本問題> 下記のデータはラットの血液中の副腎皮質ホルモン($\mu\text{g}/100\text{ml}$)を日と夜間に調べたものである。

	(対数で表示)								
日 昼	0.813	0.820	0.771	0.908	0.944	0.771	0.763	0.996	0.886
夜 間	1.045	1.100	1.013	1.009	1.204	0.996	1.352	1.173	0.934

(出典: 千葉百子氏(順天堂大学医学部衛生学教室)の御好意による)

- i) 有意水準5%で等分散の検定を行え。
- ii) 有意水準5%で平均の差を両側検定せよ。検定法はi)の結果にもとづいて選択せよ。

iii) 平均の差の95%信頼区間をつくれ。

3.2' <1元配置> 織物の樹脂加工工程で、樹脂濃度(A)が織物の引張り強さに与える影響を調べるために、それを4水準にとって実験し、下記のデータを得た。これを分析せよ。また、各水準の平均に95%信頼区間を付したグラフを作成せよ。

温 度	A_1	A_2	A_3	A_4
デ リ タ (kg)	70	74	78	76
	73	74	75	78
	75	77	79	75
	72	75	80	75

3.3 <平方和の分解>

平方和の分解(3.27)が成立することを確かめよ。

3.4 <平方和の期待値>

2元配置のモデル(3.25)から得られる

$$\bar{y}_{ij..} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \bar{\varepsilon}_{ij..}$$

$$\bar{y}_{i..} = \mu + \alpha_i + \bar{\varepsilon}_{i..}$$

$$\bar{y}_{.j..} = \mu + \beta_j + \bar{\varepsilon}_{.j..}$$

$$\bar{y}... = \mu + \bar{\varepsilon}...$$

から平方和の期待値に関する(3.31)が成立することを示せ。

3.5 <処理平均の差の推定> 2元配置で交互作用が有意なとき、 $\mu_{ij} - \mu_{i'j'}$ は $\bar{y}_{ij..} - \bar{y}_{i'j'..}$ で推定する。これに対する信頼係数 $1-\alpha$ の信頼限界をつくれ。交互作用が有意でなく、主効果が両方とも有意な場合、 $\mu_{ij} - \mu_{i'j'}(i \neq i', j \neq j')$ は $(\bar{y}_{i..} + \bar{y}_{.j..}) - (\bar{y}_{i'..} + \bar{y}_{.j'})$ で推定する。この分散が $2\sigma^2(a+b)/abr$ であることを確かめよ。これに対する信頼係数 $1-\alpha$ の両側信頼限界をつくれ。

3.6' <繰り返しのある2元配置> ある化学製品の製造工程で、収量を高める目的で反応温度(A)を3水準、触媒量(B)を4水準に選んで繰り返し $r=2$ の実験を行った結果、次のデータを得た(単位は%)。これを分析せよ。また、最適な水準組合せにおける平均を推定せよ。

	B_1	B_2	B_3	B_4
A_1	54, 55	63, 60	61, 58	51, 53
A_2	64, 67	70, 70	67, 62	58, 61
A_3	54, 50	48, 64	57, 47	47, 42

第4章 最尤法

不良率や生存率のモデル化に線形モデルはなじまない。たとえば単回帰モデルだと説明変数の値によって容易に[0, 1]の範囲を飛び出してしまうからである。また、寿命データのように正値のみをとり、右へ長く延びた分布では正規線形モデルがなじまない。近年、このような非正規分布が推測によく用いられるようになったが、その基本的道具が最尤法である。そこで、その基本事項を解説した。

4.1 一般線形モデル

線形モデルにおいては最小二乗法が統計的推測の基本的道具であった。しかしながら、線形モデルは基本的に正規分布の期待値のモデルに適切であり、それ以外の一般的な分布には必ずしもなじまない。

たとえば、毒物の量 x に対する死亡率 p のモデルとして単回帰モデル

$$(4.1) \quad p = \beta_0 + \beta_1 x$$

を想定したとする。いま、毒物の量 x_1, \dots, x_a に対し、それぞれ r 匹の動物を実験し、死亡数 y_1, \dots, y_a が観測されたとする。このとき、確率分布としては x_i ごとに独立な二項分布

$$(4.2) \quad rC_{y_i} p_i^{y_i} (1-p_i)^{r-y_i}, \quad i=1, \dots, a$$

が想定され、(4.1)のモデルを採用すると

$$E\left(\frac{y_i}{n}\right) = p_i = \beta_0 + \beta_1 x_i$$

という線形モデルを想定したことになる。ここで、最小二乗法によって $\hat{p} = \hat{\beta}_0 + \hat{\beta}_1 x$ というあてはめを行うと、 x の値によっては \hat{p} が容易に $[0, 1]$ の範囲を飛び出してしまい、これは不合理である。また

$$V\left(\frac{y_i}{n}\right) = \frac{p_i(1-p_i)}{n}$$

が x_i によるため、最小二乗法の前提である分散の一様性も成り立たない。

そこでよく用いられるのが、 p のロジット変換 logit transformation あるいは単にロジット $\log(p/(1-p))$ に線形モデル

$$(4.3) \quad \log \frac{p}{1-p} = \beta_0 + \beta_1 x$$

を仮定することである。 $p/(1-p)$ は p の単調関数で $(0, \infty)$ の値をとり、ロジットは $(-\infty, \infty)$ の値をとるので先のような不合理性は生じない。このモデルはまた次のように解釈することもできる。たとえば、各動物は毒物に対する閾値 X を持つており、投与された毒物の量 x が閾値を上回ったときに死亡する。閾値 X は動物によって異なり、母集団全体ではロジスティック分布 logistic distribution

$$(4.4) \quad P(X \leq x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

$$= \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

に従う。このとき、毒物量 x に対する死亡率 p は閾値 X が x 以下となる割合に等しいので

$$p = P(X \leq x)$$

あるいは

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x$$

となってロジットに関する線形モデルが自然に導入される(図 4.1 参照)。

なお、両端で極限値 0, 1 をとる単調な関数としては(4.4)のほかにもいろいろな分布関数を想定することができ、たとえば標準正規分布の累積分布関数

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

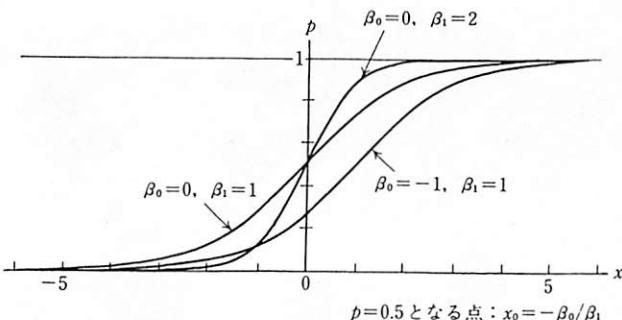


図 4.1* ロジスティック分布の分布関数

動物集団に対する毒物投与量 x と死亡率 p はこれに従う関係の一例である。 β_0, β_1 の 2 つの定数によってさまざまな位置や形をもつが、 β_0 が共通のもの、 β_1 が共通のものを比較したい。 β_0, β_1 がそれぞれ分布の中心位置、広がりを表すことが理解されよう。

を用いるのはプロビット・モデル probit model としてよく知られている。この場合、(4.3)のかわりに正規分布関数の逆関数 Φ^{-1} に関する線形モデル

$$\Phi^{-1}(p) = \frac{x - \mu}{\sigma}$$

となるが、 $\Phi^{-1}(p)$ がロジスティック分布の場合のように陽に表せないため取り扱いがやや面倒になる。また、ロジスティック分布と正規分布は前者の方がやや裾の長い分布であるが、たとえば累積確率 0.5 ± 0.3 のところで合致するように、パラメータを調節すると、表 4.1 に示すようになり、かなり広い範囲 ($p=0.1 \sim 0.9$) でよく一致する。そのため、よほど端の確率を問題にする場合以外はロジット・モデルが用いられることが多い。

ロジット・モデルは自然科学だけでなく、広く行動科学、人文科学、社会科学にも用いられる。たとえば現住所と大学の間の距離と、大学キャンパス内に住む比率のデータ(表 4.2)などにも、ロジット・モデルがよくあてはまる。

もう一つの例として表 4.3 のようなデータを考えよう。これは 1 クラス 42 名を代数と解析の期末試験成績によって二重分類した分割表である。これは 9 個の項目(セルとよばれる)を持つ多項分布で、 (i, j) セルの生起確率を p_{ij} で表すと確率分布は

$$(4.5) \quad \frac{n!}{\prod_i \prod_j y_{ij}!} \prod_i \prod_j p_{ij}^{y_{ij}}, \quad p.. = 1$$

で与えられる。ただし、 n は総観測数である。これを多項分布モデル multinomial model という。

表4.1[†] ロジット・モデルとプロビット・モデルの比較

p	ロジット [*] $0.6071 \cdot \log \{p/(1-p)\}$	プロビット $\phi^{-1}(p)$
0.001	-4.193	-3.090
0.01	-2.790	-2.330
0.05	-1.788	-1.645
0.10	-1.334	-1.285
0.20	-0.842	-0.842
0.30	-0.514	-0.524
0.40	-0.246	-0.253
0.50	0.000	0.000
0.60	0.246	0.253
0.70	0.514	0.524
0.80	0.842	0.842
0.90	1.334	1.285
0.95	1.788	1.645
0.99	2.790	2.330
0.999	4.193	3.090

*) 対数は自然対数 \log_e 。

表4.2[†] 大学への距離とキャンパス内に住む比率の関係(アメリカ)

距離(マイル)	0	10	20	30	40	50	60	70
比 率	0.108	0.258	0.500	0.742	0.892	0.960	0.986	0.995

(出典: ルビンフェルド, ビンダイク)

ロジット・モデルはこのような行動科学のデータにも用いられる。

表4.3[†] 期末試験成績のデータ(分割表)

代数 解析	優	良	可	計
優	4	2	3	9
良	8	4	6	18
可	6	3	6	15
計	18	9	15	42

(出典: 広津 1983)

ある大学の工学部のものである。この場合の興味は代数がよくできれば(できなければ)解析もよくできる(できない)といった傾向があるかないかということである。それは分割表の独立性の検定で確かめることができる(表4.5も参照)。

ここでデータ解析の興味は代数の成績と解析の成績の独立性であろう。もし、代数ができれば解析もできるという正の相関があれば表5.1において対角要素が大きくなる。一方、両者の成績が独立であれば、代数(解析)の成績のいかんにかかわらず解析(代数)の成績分布はほぼ同様となり、各行(あるいは各列)の頻度が比例的になる。このことは条件付確率が一定

$$p_{ij}/p_{i\cdot} \equiv p_{\cdot j} \quad (i \text{ に無関係})$$

あるいは

$$p_{ij}/p_{\cdot j} \equiv p_{i\cdot} \quad (j \text{ に無関係})$$

というモデルで表現できるが、これはまた、独立性の仮説

$$(4.6) \quad p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$$

とも同値である。この p_{ij} のモデルは母数の積の形で表現されるため乗法モデル multiplicative model とよばれるが、 $\log p_{ij}$ に関しては

$$\log p_{ij} = \log p_{i\cdot} + \log p_{\cdot j} = \alpha_i + \beta_j$$

のような線形モデルとなるので対数線形モデル log-linear model ともよばれる。(これについては、第II巻も参照のこと。)

(4.3) や (4.6) のように、期待値のある関数(リンク関数という)に線形モデルを仮定したものを一般線形モデル generalized linear model とよぶ。その推測に最小二乗法に代って用いられるのが最尤法である。

4.2 最尤法

(4.2) あるいは(4.5)のようないくつかの確率分布を、標本(データ)

$$\mathbf{y} = (y_1, \dots, y_n)'$$

は与えられたものとして、未知母数

$$\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$$

の関数と見たとき尤度 likelihood とよび $L(\boldsymbol{\theta})$ と表す。

ここで尤度 $L(\boldsymbol{\theta})$ を最大にする $\boldsymbol{\theta}$ の値 $\tilde{\boldsymbol{\theta}}$ は最尤推定量 maximum likelihood estimator とよばれ^{*}、現実に得られたデータで評価したときに確率最大

*) 推定量は、一般に $\hat{\boldsymbol{\theta}}$ のように $\hat{\boldsymbol{\theta}}$ (ハット)を用いるので、最小二乗推定量と区別するために最尤推定量は $\tilde{\boldsymbol{\theta}}$ (ティルダ)を用いた。

を与える推定量として、統計学上の意味をもつ

たとえば、正規線形モデルのうちもっとも簡単な繰り返し測定

$$(4.7) \quad y_i = \mu + \varepsilon_i, \quad i=1, \dots, n$$

の場合、誤差の独立性から尤度が

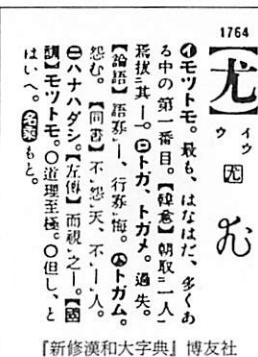
$$(4.8) \quad L(\mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(y_i - \mu)^2}{2\sigma^2} \right\}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum (y_i - \mu)^2 \right\}$$

で与えられるので、最尤推定量は、対数をとり偏微分 $\frac{\partial}{\partial \mu}, \frac{\partial}{\partial \sigma^2}$ を実行して、容易に

$$\tilde{\mu} = \bar{y}, \quad \tilde{\sigma}^2 = \frac{1}{n} \sum (y_i - \bar{y})^2$$

が得られる。この場合、 $\tilde{\mu}$ は最良不偏推定量と一致する(第 2 章参照)。 $\tilde{\sigma}^2$ は



「新修漢和大字典」博友社

を意味すると、漢和辞典にあるから、少々ややこしい。（そのほか、「けれども」「但し」の意味での「尤」もある。）

ある事実や症例がそこにあると、探偵や医師はその事実を生み出す確率が最大になるような原因を求めるだろう。つまり、その事実をもつとももっともにする原因を探し出すものである。英語ではもっともを‘likely’といい、もっともらしさをフィッシャーに従い‘likelihood’という。それを最大(maximum)にするのが最尤法である。

4.2 最 尤 法

最良な

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2$$

と分母が 1 だけ異なるが n が十分大きければそのことはあまり問題にならない。一般の正規線形モデル(2.9)の場合も θ を定める方程式は最小二乗法の正規方程式に帰着することが容易に確かめられる。さらに、4.4 節では最尤推定量が正規分布のモデル以外にも、 n が大きいときいろいろ好ましい性質をもつていることが示される。

なお、 $L(\theta)$ の最大という代りに、これと単調な関係にある $\log L(\theta)$ の最大といつてもよい。これは対数尤度 log likelihood という。

したがって、最尤推定量 $\hat{\theta}$ は

$$\frac{\partial}{\partial \theta} \log L(\theta) = 0$$

を満たす。これを尤度方程式 likelihood equation という。よく用いられる確率モデルでは、対数尤度のほうが操作が簡単になるものが多い。

一般に、ベクトル変数による微分 $\partial f(\theta)/\partial \theta$ は偏微分操作 $\partial f(\theta)/\partial \theta_i$ を行って、その結果を列ベクトルに表したものである

この対数尤度の2階微分は、母数の推定に関し、データのもつ情報量と密接に関連した量であることが示される(4.3節参照)

例 4.1 <2 項分布モデル> $Bi(n, p)$ の場合には、確率分布は

$$_nC_y p^y(1-p)^{n-y}$$

で与えられる。これを尤度と考えると図 4.2 のようになる。したがって、対数尤度

$$(4.9) \quad \log L(p) = \log \eta C_y + y \log p + (n-y) \log (1-p)$$

が得られる。 (4.9) を最大にする α は、 α について微分して

$$\tilde{p} = y/n$$

と得られる。

例 4.2 〈多項分布モデル：分割表の場合〉 (4.5) の場合には 行数 列

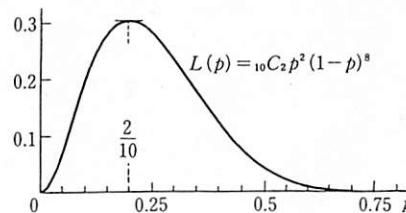


図 4.2^t 2項分布に従うデータの尤度関数($n=10, y=2$ の場合)

並んだ硬貨を机の上に投げたときそれが上を向く確率を p としよう。10回投げて2回その結果が起った。この事象の確率 $L(p)$ を p の値ごとに計算するとこここの図となるが、ちょうど $p=2/10$ のとき、この確率は最大になる。この値 0.2 をもって p の推定値とするのが、最尤推定量である。

数をそれぞれ a, b とすると対数尤度は、すべての p_{ij} の関数として

$$(4.10) \quad \log L(\mathbf{p}) = \log \left(n! / \prod_{i=1}^a \prod_{j=1}^b y_{ij}! \right) + \sum_{i=1}^a \sum_{j=1}^b y_{ij} \log p_{ij}, \quad p_{..}=1$$

となる。いま $p_{..}=1$ なので、(4.10)を最大化するのに ab 個の母数 $p_{11}, p_{12}, \dots, p_{ab}$ を自由に動かすことはできない。

そこで(4.10)から $p_{ab}=1-p_{11}-p_{12}-\dots-p_{a,b-1}$ を消去して

$$(4.11) \quad \begin{aligned} \log L(\mathbf{p}) = & \log \left(n! / \prod_i \prod_j y_{ij}! \right) + y_{11} \log p_{11} + y_{12} \log p_{12} + \dots \\ & + y_{ab} \log (1 - p_{11} - p_{12} - \dots - p_{a,b-1}) \end{aligned}$$

のように表す。(4.11)を p_{11} で偏微分して 0 とおくと

$$\frac{\partial \log L(\mathbf{p})}{\partial p_{11}} = \frac{y_{11}}{p_{11}} - \frac{y_{ab}}{1 - p_{11} - p_{12} - \dots - p_{a,b-1}} = \frac{y_{11}}{p_{11}} - \frac{y_{ab}}{p_{ab}} = 0$$

が得られる。他の p_{ij} についても同様だから

$$(4.12) \quad \frac{y_{11}}{p_{11}} = \frac{y_{12}}{p_{12}} = \dots = \frac{y_{ab}}{p_{ab}}$$

が得られ、加比の定理によって

$$\frac{y_{11}}{p_{11}} = \frac{y_{12}}{p_{12}} = \dots = \frac{y_{ab}}{p_{ab}} = \frac{y_{..}}{p_{..}} = \frac{n}{1}$$

が成り立つ。これから最尤推定量

$$(4.13) \quad \tilde{p}_{ij} = \frac{y_{ij}}{n}, \quad i=1, \dots, a, \quad j=1, \dots, b$$

が得られる。なお、これは制約付最小化問題なのでラグランジュの未定係数法で解くこともできる。その方法を独立モデル(4.6)の母数 $p_{..}, p_{..}$ の最尤推定量

を求めるために適用してみよう。

いま、モデル(4.6)の下では対数尤度が

$$(4.14) \quad \log L(\mathbf{p}) = \log \left(n! / \prod_i^a \prod_j^b y_{ij}! \right) + \sum_i^a y_{i..} \log p_{i..} + \sum_j^b y_{..j} \log p_{..j}$$

と表される。これを条件

$$\sum_i p_{i..} = 1, \quad \sum_j p_{..j} = 1$$

の下で最大化するには両式に対するラグランジュ未定係数を λ_1, λ_2 として関数

$$(4.15) \quad K = \sum_i y_{i..} \log p_{i..} + \sum_j y_{..j} \log p_{..j} + \lambda_1 (\sum_i p_{i..} - 1) + \lambda_2 (\sum_j p_{..j} - 1)$$

を最大化すればよい。 K を $p_{i..}, p_{..j}, \lambda_1, \lambda_2$ でそれぞれ偏微分して 0 とおくと次式が得られる。

$$(4.16) \quad \frac{\partial K}{\partial p_{i..}} = y_{i..}/p_{i..} + \lambda_1 = 0, \quad i=1, \dots, a$$

$$(4.17) \quad \frac{\partial K}{\partial p_{..j}} = y_{..j}/p_{..j} + \lambda_2 = 0, \quad j=1, \dots, b$$

$$(4.18) \quad \frac{\partial K}{\partial \lambda_1} = \sum_i p_{i..} - 1 = 0$$

$$(4.19) \quad \frac{\partial K}{\partial \lambda_2} = \sum_j p_{..j} - 1 = 0$$

(4.16)から $p_{i..} = -y_{i..}/\lambda_1$ が得られるので、それを(4.18)に代入することにより $\sum_i (-y_{i..}/\lambda_1) = -n/\lambda_1 = 1$ より $\lambda_1 = -n$

が得られる。したがって、最尤推定量

$$\tilde{p}_{i..} = -y_{i..}/(-n) = y_{i..}/n, \quad i=1, \dots, a$$

が得られる。まったく同様に

$$\tilde{p}_{..j} = y_{..j}/n, \quad j=1, \dots, b$$

となるので、(4.6)の下での p_{ij} の最尤推定量

$$(4.20) \quad \tilde{p}_{ij} = y_{ij}/n^2$$

が得られる。(なお、このラグランジュ未定係数法はいろいろな制約付最適化問題に応用できるので慣れておくとよい。)

実際に表 4.3 の期末試験のデータに適用すると、生起確率 p_{ij} が特別な構造を持たない場合の最尤推定値が(4.13)から表 4.4(a)のように得られる。一方、独立性のモデル(4.6)の下では(4.20)から表 4.4(b)の結果が得られる。

表4.4 表4.3の成績データへの多項分布モデルあてはめ
(a) p_{ij} が特別な構造を持たない場合

代数	解析	優	良	可	計
優		0.095	0.048	0.071	0.214
良		0.190	0.095	0.143	0.429
可		0.143	0.071	0.143	0.357
	計	0.429	0.214	0.357	1.000

(b) 独立モデル $p_{ij} \equiv p_i \cdot p_j$ の場合

代数	解析	優	良	可	計
優		0.092	0.046	0.076	0.214
良		0.184	0.092	0.153	0.429
可		0.153	0.076	0.127	0.357
	計	0.429	0.214	0.357	1.000

(a) は表4.3からふつうに相対度数を求めたものになっている。
(b) は、独立性を仮定したので、(優, 優)のセルは、 $0.214 \times 0.429 = 0.092$ 等と求められる。(1), (2) が似ていることから、代数、幾何の両成績が独立であることが示唆される。

表4.4(a)と(b)はきわめてよく似通っている。もし、データが独立モデルに適合していないと制約式(4.6)は母数を不適切な領域に制限することになり、制約なしの最大化と大きな食い違いが生じるはずである。それに対し、この例で(a), (b)がよく合致することは(4.6)が実質的な制約にはっていない、つまり、 p_{ij} が実際に独立モデル $p_i \cdot p_j$ に従っている可能性が強いことを示唆する。このことを検証する統計的方法として後で述べる尤度比検定、適合度検定を考えられる。

このように、正規分布モデルに限らず多くの場合に、最尤推定量は合理的な推定量を与えるとともにモデルの適合度に関する情報も与える。

4.3 データのもつ情報量

4.3.1 フィッシャー情報量

原点を通る单回帰モデル

$$(4.21) \quad y_i = \beta x_i + \varepsilon_i, \quad i=1, \dots, n$$

で、誤差 ε_i が正規分布 $N(0, \sigma^2)$ に従っているとする。いま、簡単のため σ^2 は既知とし、傾き β を推定する問題を考える。表現を簡潔にするため(4.21)をベクトル表示で

$$(4.22) \quad \mathbf{y} = \beta \mathbf{x} + \boldsymbol{\varepsilon}$$

と表すと、対数尤度が

$$(4.23) \quad \log L(\beta) = \log \left[(2\pi\sigma^2)^{-\frac{n}{2}} \exp \left\{ -\frac{(\mathbf{y} - \beta \mathbf{x})'(\mathbf{y} - \beta \mathbf{x})}{2\sigma^2} \right\} \right] \\ = -\frac{n}{2} \log (2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \beta \mathbf{x})'(\mathbf{y} - \beta \mathbf{x})$$

のように得られる。これを最大にする β は線形模型(4.22)に対する最小二乗解にほかならない。また第2章の結果から最尤推定量が

$$(4.24) \quad \tilde{\beta} = \hat{\beta} = \mathbf{x}' \mathbf{y} / \mathbf{x}' \mathbf{x}$$

のように得られる。(4.23)は β に関する2次式だから図4.3(b)のように表され、 $\tilde{\beta}$ はその頂点の横座標である。

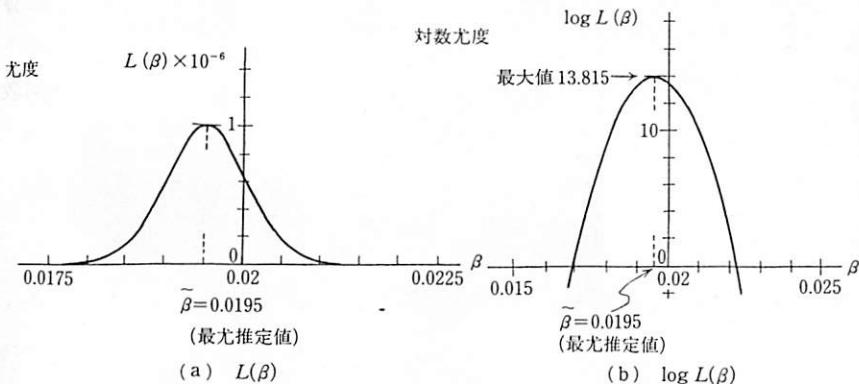


図4.3 原点を通る单回帰モデルにおける傾き β の尤度、対数尤度

データは、下記のごとく機械加工部品の焼入れ時間 x (秒)、焼入れ深さ y (インチ)である。一見して比例関係が見受けられるが、傾き β の最尤推定値 0.0195 が得られる。また、この図のピーク(頂点)の鋭さも重要である。

x	5	10	15	20	25	30	(秒)
y	0.12	0.23	0.29	0.37	0.46	0.61	(インチ)

(出典: ボウカーリーバーマン)

最大値: $\tilde{\beta} = \sum x_i y_i / \sum x_i^2 = 44.45 / 2275 = 0.0195$ が得られる。その時の最大が $\max \log L = 13,815$, $\max L \neq 10^6$ のように得られる。

ここで $\log L(\beta)$ の 2 階微分

$$(4.25) \quad \frac{\partial^2 \log L(\beta)}{\partial \beta^2} = -\mathbf{x}' \mathbf{x} / \sigma^2$$

は β^2 の係数にほかならず、頂点での(この例では(4.25)が β によらないので頂点に限らないが)曲率に関する量である。符号一は上に凸であることを表す。ここで $\mathbf{x}' \mathbf{x} / \sigma^2$ が大きければ図 4.3(b)の頂点は急峻で、 $\tilde{\beta}$ から少しづれただけで $\log L(\beta)$ の値は急激に減少し、一方 $\mathbf{x}' \mathbf{x} / \sigma^2$ が小さければ頂点は偏平で、広い範囲の β が $\tilde{\beta}$ とあまり変わらない $\log L(\beta)$ の値を与える。このことは $\tilde{\beta}$ の推定量としての良さに対応し、実際この例で

$$(4.26) \quad (\mathbf{x}' \mathbf{x} / \sigma^2)^{-1} = (\mathbf{x}' \mathbf{x})^{-1} \sigma^2$$

は $\tilde{\beta}$ の分散にほかならない。分散は小さければ小さいほど正確さをあらわすという意味で $\mathbf{x}' \mathbf{x} / \sigma^2$ はデータ \mathbf{y} の持つ情報量を表しており、最初にこのような概念を明確にしたフィッシャーの名を冠して、とくにフィッシャー情報量 Fisher information とよばれる。たとえば実験点 \mathbf{x} を実験者が選べる場合には、可能なかぎり $\mathbf{x}' \mathbf{x}$ を大きくするとフィッシャー情報量が増え推定量の分散が小さくなる。

一般には、対数尤度の 2 階微分はデータ \mathbf{y} に依存するので、 \mathbf{y} が密度関数 $f_\theta(\mathbf{y})$ に従っている^{*)}ときに期待値をとると

$$(4.27) \quad \begin{aligned} I(\theta) &= -E\left\{ \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right\} \\ &= -E\left\{ \frac{\partial^2 \log f_\theta(\mathbf{y})}{\partial \theta^2} \right\} \end{aligned}$$

となる。これをフィッシャー情報量とよぶ。さらにベクトル母数 $\theta = (\theta_1, \dots, \theta_p)'$ のときは

$$\begin{aligned} I_{ij}(\theta) &= -E\left\{ \frac{\partial^2 \log L(\theta)}{\partial \theta_i \partial \theta_j} \right\} \\ &= -E\left\{ \frac{\partial^2 \log f_\theta(\mathbf{y})}{\partial \theta_i \partial \theta_j} \right\} \end{aligned}$$

を (i, j) 要素とする $p \times p$ 行列 $I(\theta)$ を、フィッシャー情報行列 Fisher's information matrix とよぶ。

^{*)} $f(\mathbf{y}, \theta)$ と書くことも多いが、この巻では θ は $f(\mathbf{y})$ の添字とした。

統計的推定におけるフィッシャー情報量の意義は、次に述べる不偏推定量の分散に関するクラメール・ラオの下限でいっそう明解になる。

4.3.2 クラメール・ラオの下限

データ \mathbf{y} が密度関数 $f_\theta(\mathbf{y})$ に従っているとして 1 次元の未知パラメータ θ の推定を考えよう。ここで推定量 $t(\mathbf{y})$ が必ずしも不偏でなくてよいとするならいくらでも分散の小さな推定量を作ることができる。たとえばデータ \mathbf{y} とは無関係に定数 1 を推定量とする

$$t(\mathbf{y}) \equiv 1$$

は分散 0 であり、真の値 θ がたまたま 1 ならこの上なくよい推定量である。しかし、 θ は実際にどのような値であるか分からぬのだから、このような分散 0 の定数推定量は無意味である。そこで θ がどのような値であろうとそれを不偏に推定したいものとすると、おのずと分散はある程度より小さくはできないことが想像できる。後の(4.38)の不等式で見るよう、その下限すなわち不偏推定量でそれ以上分散を小さくできないという分散の限界を与えるのが、これから述べるクラメール・ラオの下限^{*)}である。以下にそれを導こう。

統計量 $t(\mathbf{y})$ が θ の不偏推定量であるということは

$$(4.28) \quad \int t(\mathbf{y}) f_\theta(\mathbf{y}) d\mathbf{y} = \theta \quad (\mathbf{y} = (y_1, \dots, y_n)', \quad d\mathbf{y} \equiv dy_1 dy_2 \cdots dy_n)$$

と表される(ここで、積分は n 次元空間全域にわたる定積分である)。この式の両辺を θ で微分することにより

$$(4.29) \quad \int t(\mathbf{y}) \{ \partial f_\theta(\mathbf{y}) / \partial \theta \} d\mathbf{y} = 1$$

が得られる。なお、左辺で微分を積分記号の中で行っているが、以下ではとくに断わらないが、密度関数 $f_\theta(\mathbf{y})$ はこのように微分・積分の順序が交換できるなどの正則条件は満たしているものとする。一方、全域での $f_\theta(\mathbf{y})$ の積分

^{*)} クラメール(H. Cramér)は今世紀の統計学者で、代数学者のクラメール(G. Cramer, 1704-1752)とは全く別人である。

(4.30)

$$\int f_\theta(\mathbf{y}) d\mathbf{y} \equiv 1$$

の両辺を θ で微分することにより

(4.31)

$$\int \{\partial f_\theta(\mathbf{y})/\partial \theta\} d\mathbf{y} = 0$$

が得られるから、(4.29)に定数 θ をつけ加えて

(4.32)

$$\int \{t(\mathbf{y}) - \theta\} \{\partial f_\theta(\mathbf{y})/\partial \theta\} d\mathbf{y} = 1$$

とすることができる。(4.32)は

$$\begin{aligned} & \int \{t(\mathbf{y}) - \theta\} \{\partial \log f_\theta(\mathbf{y})/\partial \theta\} f_\theta(\mathbf{y}) d\mathbf{y} \\ &= E[\{t(\mathbf{y}) - \theta\} \cdot \{\partial \log f_\theta(\mathbf{y})/\partial \theta\}] = 1 \end{aligned}$$

と表され、これは $g(\mathbf{y}) = t(\mathbf{y}) - \theta$ と $h(\mathbf{y}) = \partial \log f_\theta(\mathbf{y})/\partial \theta$ の共分散が 1 であることを意味する。ところが一般に、相関係数の絶対値は 1 以下であるから

$$\frac{\{\text{Cov}(g, h)\}^2}{V(g) \cdot V(h)} = \frac{1}{V\{t(\mathbf{y})\} \cdot V\{\partial \log f_\theta(\mathbf{y})/\partial \theta\}} \leq 1$$

が成り立つ。したがって、不偏推定量 $t(\mathbf{y})$ の分散に関し、不等式

(4.33)

$$V\{t(\mathbf{y})\} \geq 1/V\{\partial \log f_\theta(\mathbf{y})/\partial \theta\}$$

が成り立つ。

これで不偏推定量の分散には、これ以上は小さくできないという下の限界があることがわかった。目的の大半は達せられたが、なお右辺を見やすくすることができる。(4.31)により

(4.34)

$$E(\partial \log f_\theta(\mathbf{y})/\partial \theta) = 0$$

が成り立つことに注意すると、(4.33)の右辺の分母が

(4.35)

$$V\{\partial \log f_\theta(\mathbf{y})/\partial \theta\} = E\{\partial \log f_\theta(\mathbf{y})/\partial \theta\}^2$$

となることがわかる。一方、

(4.36)

$$\begin{aligned} \frac{\partial^2 \log f_\theta(\mathbf{y})}{\partial \theta^2} &= \frac{\partial}{\partial \theta} \left\{ \frac{\partial f_\theta(\mathbf{y})/\partial \theta}{f_\theta(\mathbf{y})} \right\} \\ &= \frac{\partial^2 f_\theta(\mathbf{y})}{\partial \theta^2} \cdot \frac{1}{f_\theta(\mathbf{y})} - \left\{ \frac{\partial f_\theta(\mathbf{y})/\partial \theta}{f_\theta(\mathbf{y})} \right\}^2 \end{aligned}$$

の期待値をとると、右辺第 1 項は

$$\int \frac{\partial^2 f_\theta(\mathbf{y})}{\partial \theta^2} \cdot \frac{1}{f_\theta(\mathbf{y})} \cdot f_\theta(\mathbf{y}) d\mathbf{y} = \int \frac{\partial^2 f_\theta(\mathbf{y})}{\partial \theta^2} d\mathbf{y}$$

となる。これは(4.31)の両辺を θ で微分したものだから 0 である。よって、右辺第二項から、等式

$$(4.37) \quad -E\left\{ \frac{\partial^2 \log f_\theta(\mathbf{y})}{\partial \theta^2} \right\} = E\left\{ \frac{\partial \log f_\theta(\mathbf{y})}{\partial \theta} \right\}^2$$

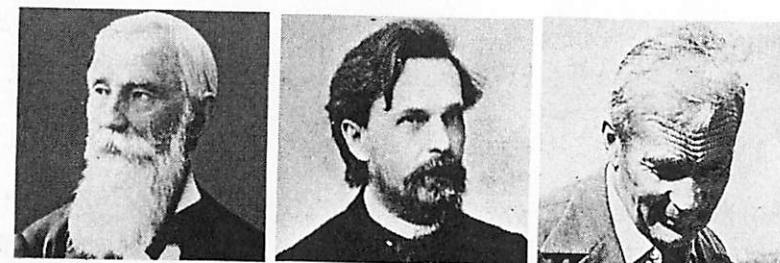
が出る。これが(4.35)、つまりは不等式(4.33)の右辺の分母を与えていた。これはまた(4.27)のフィッシャー情報量 $I(\theta)$ の別表現にほかならない。以上により、不偏推定量の分散についての不等式(4.33)はフィッシャー情報量を用い

※ロシア人と統計学※

ガウスがどういう学問上の動機か、年とってからロシア語に関心を示し、思い立って手習いを始めたことは、割に知られていない。統計学はその生い立ちから「イギリス人の学問」という感じもあるが、ロシア人の理論的貢献も多い。まずはチェビシェフ、マルコフ、コルモゴロフのトリオがあげられる。

ところで、有名な「ガウス・マルコフの定理」は、最初はマルコフによると考えられていたが、これは誤りでガウスが既に証明しているとわかり、ドイツとロシアで名前を分け合ってこの名前となった。ロシア人の気持を考えてのことだろう。

ついでながら、ラオはインドの統計学者であるが、インドもマハラーナビスをはじめとして多くの統計学者を世界に送り出している。



P. L. チェビシェフ
(1821-1894)

チエビシェフの不等式、大数の法則で有名。

A. A. マルコフ
(1856-1922)

ガウス・マルコフの定理、マルコフ連鎖の創始者。

A. N. コルモゴロフ
(1903-)

公理論的確率論の基礎を作った。

て

$$(4.38) \quad V\{t(y)\} \geq \frac{1}{I(\theta)}$$

と表される。これをクラメール・ラオの不等式 Cramér-Rao's inequality, また右辺をクラメール・ラオの下限 Cramér-Rao's lower bound という。すなわち、フィッシャー情報量の逆数がクラメール・ラオの下限を与える。これで統計的推定におけるフィッシャー情報量の意味が明確になった。

ここで、多くの場合そうであるように、データ y_1, \dots, y_n が独立であるとすると

$$\begin{aligned} \log f_\theta(y) &= \log \{f_\theta(y_1) \cdots f_\theta(y_n)\} \\ &= \sum_{i=1}^n \log f_\theta(y_i) \end{aligned}$$

と表され、 $f_\theta(y_i)$ がまた独立、同一分布に従うことから

$$\begin{aligned} (4.39) \quad I(\theta) &= -E\left\{\sum_{i=1}^n \frac{\partial^2 \log f_\theta(y_i)}{\partial \theta^2}\right\} \\ &= -nE\left\{\frac{\partial^2 \log f_\theta(y)}{\partial \theta^2}\right\} \\ &= nI_1(\theta) \end{aligned}$$

と表される。ただし、 $I_1(\theta)$ はデータ 1 個当たりのフィッシャー情報量である。この場合は、クラメール・ラオの下限は $1/\{nI_1(\theta)\}$ で、 n に反比例する。

クラメール・ラオの下限を達成する不偏推定量を有効推定量 efficient estimator という。有効推定量は、その定義から最小分散不偏推定であるが、その逆は必ずしも成りたたない。たとえば、繰り返し測定のモデル(4.7)で

$$\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)$$

は最小分散不偏推定量であるが、その分散 $2\sigma^4/(n-1)$ はクラメール・ラオの下限 $2\sigma^4/n$ より大きく、有効推定量とはならない。以下の 2 例は有効推定量の例である。

例 4.3 〈繰り返し測定のモデル〉 (4.7) の μ の最尤推定は $\tilde{\mu} = \bar{y}$ とわかっている。 \bar{y} は不偏でその分散は $V(\bar{y}) = \sigma^2/n$ である。一方、

$$\begin{aligned} (4.40) \quad I_1(\mu) &= -E\left[\frac{\partial^2}{\partial \mu^2} \log \left\{\frac{1}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2}\right\}\right] \\ &= E\left[\frac{\partial^2}{\partial \mu^2} \left\{\frac{(y-\mu)^2}{2\sigma^2}\right\}\right] \\ &= 1/\sigma^2 \end{aligned}$$

だからクラメール・ラオの下限は

$$1/\{n(1/\sigma^2)\} = \sigma^2/n$$

である。 \bar{y} の分散はこの下限を達成しており、 \bar{y} が有効推定量であることがわかる。

なお、(4.37) の後の表現からも

$$\begin{aligned} I_1(\mu) &= E\left[\frac{\partial}{\partial \mu} \left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}\right]^2 = E\left\{\left(\frac{y-\mu}{\sigma^2}\right)^2\right\} \\ &= E(y-\mu)^2/\sigma^4 = \sigma^2/\sigma_4 = 1/\sigma^2 \end{aligned}$$

となって、確かに(4.40)と同じ結果が得られる。

例 4.4 〈ポアソン分布〉 y_1, \dots, y_n が互いに独立にポアソン分布

$$(4.41) \quad f_\lambda(y) = \frac{e^{-\lambda} \lambda^y}{y!}, \quad y=0, 1, 2, \dots$$

に従う場合

$$\begin{aligned} (4.42) \quad I_1(\lambda) &= -E\left\{\frac{\partial^2}{\partial \lambda^2} \left(\log \frac{e^{-\lambda} \lambda^y}{y!}\right)\right\} \\ &= -E\left\{\frac{\partial^2}{\partial \lambda^2} (-\lambda + y \log \lambda)\right\} \\ &= -E\left\{\frac{\partial}{\partial \lambda} \left(-1 + \frac{y}{\lambda}\right)\right\} \\ &= E\left(\frac{y}{\lambda^2}\right) \\ &= \frac{1}{\lambda} \end{aligned}$$

である。したがって、クラメール・ラオの下限は

$$1/\{n(1/\lambda)\} = \lambda/n$$

である。

他方、 $t(y) = \bar{y}$ は $E(\bar{y}) = \lambda$ から不偏であり、さらに

$$V(\bar{y}) = \frac{1}{n^2} V(y_1 + \dots + y_n) = \frac{1}{n^2} \cdot n\lambda = \frac{\lambda}{n}$$

だから、分散の下限を達成し有効推定量である。この場合も(4.37)の後の表現を使うと

$$\begin{aligned} (4.43) \quad I_1(\lambda) &= E\left\{\frac{\partial}{\partial \lambda}(-\lambda + y \log \lambda)\right\}^2 = E\left(-1 + \frac{y}{\lambda}\right)^2 \\ &= E(y - \lambda)^2 / \lambda^2 = \lambda / \lambda^2 = 1/\lambda \end{aligned}$$

となって(4.42)と同じ結果が得られる。

このようにして、われわれは、フィッシャー情報量、クラメール・ラオの下限、有効推定量の定義など、推定量の良さを判断することができるいくつかの基準を手にした。そこで、以上の準備から最尤推定量が最適であることを次のように述べることができる(最尤推定量の最適性)。すなわち、最尤推定量は漸近的に不偏な推定量(一致推定量という)であり、その漸近分散はクラメール・ラオの下限に一致する。すなわち、漸近的に有効な推定量である。さらに、その漸近分布は正規分布である。次節でこれらのことと簡単に示そう。

4.4 最尤推定量の最適性

4.4.1 一致性

簡単のため密度関数 $f_\theta(y)$ がただ一つの未知母数 θ で定められている場合を考える。このとき、 n 個の独立なデータ $\mathbf{y} = (y_1, \dots, y_n)'$ から定めた最尤推定量 $\tilde{\theta}_n$ は n 大のとき、密度関数の緩い条件(正則条件)の下で真の値 θ_0 に確率収束する。

すなわち、任意の $\epsilon > 0$ に対し

$$\lim_{n \rightarrow \infty} P\{|\tilde{\theta}_n - \theta_0| > \epsilon\} = 0 \quad (\text{確率収束})$$

が成立つ。最尤推定量のこの性質を一致性 consistency という。

[略証] $\hat{\theta}_n$ の標本の大きさを表す添字 n は自明なので省略する。いま、対数尤度が

$$\log L(\theta) = \sum_{i=1}^n \log f_\theta(y_i)$$

4.4 最尤推定量の最適性

と表されているので、最尤推定量 $\tilde{\theta}$ は、尤度方程式

$$(4.44) \quad \frac{1}{n} \cdot \frac{\partial \log L(\theta)}{\partial \theta} \equiv \frac{1}{n} \sum_{i=1}^n \frac{\partial \log f_\theta(y_i)}{\partial \theta} = 0$$

の解である($1/n$ は後の便宜のため)。上式の中央は、 n 個の同一分布に従う独立な確率変数 $\partial \log f_\theta(y_i) / \partial \theta$ ($i=1, 2, \dots, n$) の平均値なので、大数の法則により

$$(4.45) \quad E_{\theta_0} \left\{ \frac{\partial \log f_\theta(y)}{\partial \theta} \right\}$$

に確率収束する(ここで $E_\theta(\cdot)$ は $\int \cdot f_\theta(y) dy$ を意味する)。この極限値は定数だが、 θ の関数のみならず θ_0 にも依っていることを表すため、 $\mu_{\theta_0}(\theta)$ と記しておこう。ここで、 $\theta = \theta_0$ とおいた $\mu_{\theta_0}(\theta_0)$ は

$$\int_{-\infty}^{\infty} \frac{\partial f_\theta(y)}{\partial \theta} \cdot \frac{1}{f_\theta(y)} \Big|_{\theta=\theta_0} f_{\theta_0}(y) dy = \int_{-\infty}^{\infty} \frac{\partial f_\theta(y)}{\partial \theta} \Big|_{\theta=\theta_0} dy$$

で、(4.30), (4.31) で示したように 0 となる。すなわち θ_0 は方程式 $\mu_{\theta_0}(\theta) = 0$ の一つの解であるが、もしこの解がただ一つの解なら、各 n に対する尤度方程式(4.44)の解 $\tilde{\theta}_n$ も $n \rightarrow \infty$ のとき収束して

$$\tilde{\theta}_n \longrightarrow \theta_0 \quad (\text{確率収束})$$

とならなければならない。

4.4.2 漸近有効性

さて、最尤推定量は n を大きくすると真の値に近づくことがひとまずわかった。次の問題は真の値のまわりでどの位のばらつきを持つ推定量なのかを知ることである。さらに、区間推定や検定の問題を考えるときには期待値、分散だけでなく標本分布も知る必要がある。

これらについては、 y_1, \dots, y_n がたがいに独立に密度関数 $f_\theta(y)$ に従っているときは最尤推定量の定義式

$$(4.46) \quad \left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\tilde{\theta}} \equiv \sum_{i=1}^n \left. \frac{\partial \log f_\theta(y_i)}{\partial \theta} \right|_{\theta=\tilde{\theta}} = 0$$

をテイラー展開して($1/n$)より高次の項を無視することにより、次のように明解に答えることができる。

(4.46) は $\tilde{\theta}$ の関数である(実際には、値は 0). これを $\tilde{\theta} = \theta_0$ のまわりで展開することにより式

$$(4.47) \quad \sum_{i=1}^n \frac{\partial \log f_\theta(y_i)}{\partial \theta} \Big|_{\theta=\theta_0} + (\tilde{\theta} - \theta_0) \sum_{i=1}^n \frac{\partial^2 \log f_\theta(y_i)}{\partial \theta^2} \Big|_{\theta=\theta^*} = 0$$

が得られる. ただし、 θ^* は $\tilde{\theta}$ と θ_0 の間の値である($(\theta^* - \tilde{\theta})(\theta^* - \theta_0) < 0$ を満たす)が、 n を大きくすると $\tilde{\theta}$ が θ_0 に一致するため θ^* も θ_0 に一致するので、最尤推定量の漸近的性質を調べるときには、 θ^* は θ_0 に等しいと思ってよい. ここで $y_i (i=1, 2, \dots, n)$ が独立に同一の分布に従うことから、量

$$(4.48) \quad z_i = \frac{\partial \log f_\theta(y_i)}{\partial \theta} \Big|_{\theta=\theta_0}, \quad (i=1, 2, \dots, n)$$

$$(4.49) \quad w_i = \frac{\partial^2 \log f_\theta(y_i)}{\partial \theta^2} \Big|_{\theta=\theta^*}, \quad (i=1, 2, \dots, n)$$

もそれぞれ独立同一分布に従う. そこで(4.47) を \sqrt{n} で割って

$$(4.50) \quad \frac{1}{\sqrt{n}} \sum_{i=1}^n z_i + \sqrt{n}(\tilde{\theta} - \theta_0) \cdot \frac{1}{n} \sum_{i=1}^n w_i = 0$$

と表現すると、 n が大きいとき

i) $\frac{z_1 + \dots + z_n}{\sqrt{n}}$ は中心極限定理により漸近的に正規分布に従い、

ii) $\frac{w_1 + \dots + w_n}{n}$ は大数の法則によりその期待値に確率収束する.

まず、 z_i の期待値は既に(4.34) あるいは(4.45) に関連して繰り返し述べたように 0 である. 分散も(4.37) に示したように

$$(4.51) \quad V(z_i) = E_{\theta_0} \left\{ \frac{\partial \log f_\theta(y)}{\partial \theta} \Big|_{\theta=\theta_0} \right\}^2 = I_1(\theta_0)$$

と得られる. 一方、 w_i では、 $\theta^* \neq \theta_0$ とするとやはり(4.37) のもう一方より

$$(4.52) \quad E_{\theta_0}(w_i) = E_{\theta_0} \left\{ \frac{\partial^2 \log f_\theta(y)}{\partial \theta^2} \Big|_{\theta=\theta_0} \right\} = -I_1(\theta_0)$$

が得られる.

したがって、 n が大きいとき、これらを(4.50) に代入して整理することにより、 u を正規分布 $N(0, I_1(\theta_0))$ に従う確率変数として

$$\sqrt{n}(\tilde{\theta} - \theta_0) \sim \frac{u}{I_1(\theta_0)} \quad (\text{分布が同一})$$

という表現が得られる. $u/I_1(\theta_0)$ の分布は、 u に対し分散が $1/(I_1(\theta_0))^2$ 倍されるから、結局、 $\sqrt{n}(\tilde{\theta} - \theta_0)$ の漸近分布は、正規分布 $N(0, 1/I_1(\theta_0))$ であることがわかる. わかりやすくいえば、最尤推定量 $\tilde{\theta}$ は θ_0 のまわりにばらつき、 n が十分大のときその漸近分散は $1/\{nI_1(\theta_0)\}$ 、漸近分布は正規分布

$$N(\theta_0, 1/nI_1(\theta_0))$$

であるという結果が得られる. このことは $\tilde{\theta}$ が θ_0 のまわりにばらつくオーダーが $n^{-\frac{1}{2}}$ であることを意味する. 以上が、最尤推定量の漸近分布として知られているものである.

漸近分散 $1/\{nI_1(\theta_0)\}$ はちょうどクラメール・ラオの下限に等しいから、これから最尤推定量は漸近的に有効 asymptotically efficient な推定量であることがわかる. なお、 $1/\{nI_1(\theta_0)\}$ は n 大のとき 0 になるので、正確には、 $\sqrt{n}\tilde{\theta}$ の漸近分散が $1/I_1(\theta_0)$ というのが正しい表現であることに注意しよう.

例 4.5 〈ポアソン分布に対する最尤推定：漸近分布〉 データ y_1, y_2, \dots, y_n がたがいに独立に母数 λ のポアソン分布 $Po(\lambda)$ に従っているとき、尤度、対数尤度はそれぞれ

$$(4.53) \quad L(\lambda) = \prod_{i=1}^n (e^{-\lambda} \lambda^{y_i} / y_i!),$$

$$\log L(\lambda) = C - n\lambda + (\sum_i y_i) \cdot \log \lambda \quad (C \text{ は } \lambda \text{ を含まない定数})$$

であるから、尤度方程式

$$\partial \log L(\lambda) / \partial \lambda = -n + \sum_i y_i / \lambda = 0$$

を解いて、ただちに最尤推定量

$$\tilde{\lambda} = \sum_i y_i / n = \bar{y}$$

を得る. すでに例 4.4 で見たように、クラメール・ラオの下限は λ/n である.

したがって、真の値が λ_0 のとき、 $\tilde{\lambda}$ は n が大きいとき平均 λ_0 、漸近分散 λ_0/n の正規分布 $N(\lambda_0, \lambda_0/n)$ に従う.

この場合の最尤推定量 $\tilde{\lambda}$ はその分散 λ/n が、正確に ($n \rightarrow \infty$ のときはもとより、有限の n でも) クラメール・ラオの下限と一致するので、漸近的でなく小標

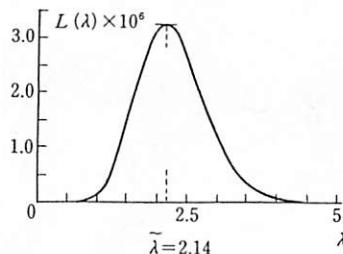


図 4.4 ポアソン分布に対する尤度関数
この例はある地方都市の 1 週間の事故死傷者数である(架空例).

日	月	火	水	木	金	土	計
5	2	0	1	3	1	3	15(n)

尤度関数は

$L(\lambda) = e^{-\lambda} \lambda^{15} / (5! 2! 0! 1! 3! 1! 3!) \quad ((4.53) \text{ で } n=7 \text{ の場合})$
となるが、それが本圖である(10^6 倍に拡大)。最大は $\lambda = 15/7 = 2.14$ とされ、これが最尤推定値を与える。このピーク(頂点)の鋭さは n に関係する。

本でも有効推定量となっている。

ここまではデータ y_1, y_2, \dots, y_n が独立で、同一分布に従うとして、最尤推定量の持つ最適性を示した。実は、これらの最適性はもっとも緩い条件でも成り立つ。たとえば、相関があっても、情報量 $I(\theta)$ が n とともに増えればよい。

4.4.3 多次元の母数*)

1 次元パラメータの場合と本質的なところは同じなので結果だけを述べる。

データ $\mathbf{y} = (y_1, y_2, \dots, y_n)'$ が p 次元の母数 $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$ によって定められる密度関数 $f_{\boldsymbol{\theta}}(\mathbf{y})$ に従っているとする。いま、データ \mathbf{y} に対するフィッシャーの情報行列を

$$(4.54) \quad I(\boldsymbol{\theta}) = -E\left\{\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\mathbf{y})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right\}$$

$$(4.55) \quad = E\left(\left\{\frac{\partial \log f_{\boldsymbol{\theta}}(\mathbf{y})}{\partial \boldsymbol{\theta}}\right\} \cdot \left\{\frac{\partial \log f_{\boldsymbol{\theta}}(\mathbf{y})}{\partial \boldsymbol{\theta}}\right\}'\right)$$

と表すと、適当な正則条件の下で、最尤推定量 $\tilde{\boldsymbol{\theta}}$ が $\boldsymbol{\theta}$ の漸近的に有効な推定量であること、および $\boldsymbol{\theta}$ の真の値が $\boldsymbol{\theta}_0$ であるとき $\sqrt{n}(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ の漸近分布は、データ 1 個当たりの情報行列

$$I_1(\boldsymbol{\theta}) = I(\boldsymbol{\theta})/n$$

によって多(か)次元正規分布 $N_p(\mathbf{0}, (I_1(\boldsymbol{\theta}))^{-1})$ となることなどが示される($^{-1}$ は逆行列を表す)。

ただし、 $\partial/\partial \boldsymbol{\theta}$ は f に対して $\partial f/\partial \theta_i$ を配置した列ベクトル、また、 $\partial^2/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'$ は同じく $\partial^2 f/\partial \theta_i \partial \theta_j$ を配置した行列を表す。期待値の操作 E は要素ごとに行う。

なお、(4.54)、(4.55)が同値であることは、(4.37)の両辺が等しいことの証明と全く同じようにして示すことができる。

※合いすぎて困った(?)※

メンデルが「メンデルの法則」の裏付けとして出した有名なエンドウ豆の 4 分類データの一つは次のものである。

表現型	黄色・丸い	黄色・しわがある	緑色・丸い	緑色・しわがある	計
観測度数 f_i	315	101	108	32	556(n)
確率 p_i	9/16	3/16	3/16	1/16	1
理論度数 np_i	312.75	104.25	104.25	34.75	556
両度数の差	2.25	-3.75	3.75	-2.25	0



観測度数 f_1, f_2, f_3, f_4 を最尤法による理論度数 np_1, np_2, np_3, np_4 と比較してわかるることは、一見して一致が相当によい、むしろ、合いすぎるとあるということである。 n が 556 とかなり大きいから漸近理論が使えて、適合度の χ^2 統計量を計算してみると

$$\chi^2 = 0.470$$

であり、当然のことながら相当に値は小さい。 $\chi^2_{0.05}(3) = 7.815$ であるから、データはメンデル理論に適合しているといえる((4.71)参照)。

ところで、いくら n が大きいとはいえ、合いすぎているのではないかという疑問も出る。 χ^2 分布の小さい方(左側)の値の限界として $\chi^2_{0.90}(3) = 0.584$ をとると、上の χ^2 の値はこれよりもまだ小さい。さる高名な統計学者は、メンデルは合わせたのではないかとさえ疑っている。ただし、その証拠はない。それにメンデルの法則は今や不動のものとして確立しているのだが……。

*) 重要ではあるが、やや進んだ内容である。

例 4.6 <正規分布に対する最尤推定：漸近分布> y_1, y_2, \dots, y_n がたがいに独立に正規分布に従っているとき、対数尤度 $\log L(\mu, \sigma^2)$ と最尤推定量 $\tilde{\mu}$, $\tilde{\sigma}^2$ はすでに、(4.8)から

$$\tilde{\mu} = \bar{y}, \quad \tilde{\sigma}^2 = \sum_i (y_i - \bar{y})^2 / n$$

と求められる。

フィッシャーの情報行列を求めてみよう。いま、(4.8)で $n=1$, $y_1 \equiv y$ とすると

$$\frac{\partial \log L(\mu, \sigma^2)}{\partial \mu} = \frac{y - \mu}{\sigma^2},$$

$$\frac{\partial \log L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{(y - \mu)^2}{2\sigma^4}$$

となる。これらのベクトルから、データ 1 個当たりの情報量は

$$\begin{aligned} I_1(\mu, \sigma^2) &= E \left(\begin{pmatrix} \frac{y - \mu}{\sigma^2} \\ \frac{(y - \mu)^2 - \sigma^2}{2\sigma^4} \end{pmatrix} \begin{pmatrix} \frac{y - \mu}{\sigma^2} & \frac{(y - \mu)^2 - \sigma^2}{2\sigma^4} \end{pmatrix} \right) \\ &= \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(2\sigma^4) \end{pmatrix} \end{aligned}$$

となる。よって、 μ_0, σ_0^2 が真の値のとき

$$\sqrt{n} \begin{pmatrix} \tilde{\mu} - \mu_0 \\ \tilde{\sigma}^2 - \sigma_0^2 \end{pmatrix} = \sqrt{n} \begin{pmatrix} \bar{y} - \mu_0 \\ \sum (y_i - \bar{y})^2 / n - \sigma_0^2 \end{pmatrix}$$

の漸近分布は、2 次元正規分布

$$N_2 \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & 1/(2\sigma_0^4) \end{pmatrix}$$

となることがわかる。なお、 $\tilde{\mu}$ の分布は小標本でも成り立つ正確な分布であり、ここにみられる $\tilde{\mu}$ と $\tilde{\sigma}^2$ が独立という性質も小標本で成立している。

正確な exact とは、漸近的 asymptotic が $n \rightarrow \infty$ の極限(近似)を考えるのに対し、有限の(止っている) n に対し、という意味である。

4.5 検定の漸近論^{*)}

4.5.1 最尤推定量に基づく検定

まず 1 次元母数の場合を考え、データ $y = (y_1, \dots, y_n)'$ の尤度が

$$L(\theta) = f_\theta(y)$$

で与えられるとする。このとき

$$(4.56) \quad \text{帰無仮説} \quad H_0: \theta = \theta_0$$

を

$$(4.57) \quad \text{対立仮説} \quad H_1: \theta = \theta_1$$

に対して検定する有意水準 α の検定で検出力が最大となる棄却域はネイマン・ピアソンの基本定理(第6章)に基づいて作れるが、分布論が複雑になり、必ずしも簡単に正確な棄却域を作れないことがある。そのようなとき、それに代って最尤推定量の漸近分布に基づいて検定を構成することができる。

そこで、具体的に、帰無仮説 H_0 の下で $\sqrt{n}(\tilde{\theta} - \theta_0)$ が漸近的に正規分布 $N(0, 1/I_1(\theta_0))$ に従うことを利用すると、 $\theta_1 > \theta_0$ の場合に棄却域

$$(4.58) \quad \sqrt{n}I_1(\theta_0)(\tilde{\theta} - \theta_0) > z_\alpha$$

が得られる。ただし、 $I_1(\theta) = I(\theta)/n$ はデータ 1 個当たりのフィッシャー情報量である。ここで z_α は標準正規分布の上側の確率が α となる点である。

例 4.7 <正規分布に対する平均の検定> y_1, \dots, y_n がたがいに独立に正規分布 $N(\mu, \sigma^2)$ に従っているとする。ただし σ^2 は既知とする。このとき

$$H_0: \mu = \mu_0 \quad H_1: \mu = \mu_1 \quad (> \mu_0)$$

の検定を最尤推定量 $\tilde{\mu}$ の漸近分布から求めよう。

この場合 $\sqrt{n}(\tilde{\mu} - \mu_0)$ の H_0 の下での漸近分布は正確な分布 $N(0, \sigma_0^2)$ だから棄却域

$$\sqrt{n}(\tilde{\mu} - \mu_0)/\sigma_0 > z_\alpha$$

が得られる。

これは $\mu_1 (> \mu_0)$ の具体的な値によらず常に最大の検出力を与える検定(一様最強力検定)である。

*) 検出力(本シリーズ I『統計学入門』第12章参照)に対する理解が必要である。本巻第6章読了後ならなおよい。

例4.8 〈ポアソン分布に対する検定〉 y_1, \dots, y_n がたがいに独立に $P_\theta(\lambda)$ に従う場合に $H_0: \lambda = \lambda_0$ の検定を最尤推定量の漸近分布に基づいて求めよう。

最尤推定量は $\tilde{\lambda} = \bar{y}$ であり、 $\sqrt{n}(\lambda - \lambda_0)$ の H_0 の下での漸近分布は $N(0, \lambda_0)$ である。したがって、 $H_1: \lambda = \lambda_1 (> \lambda_0)$ の場合に棄却域

$$\sqrt{n}(\bar{y} - \lambda_0) / \sqrt{\lambda_0} > z_\alpha$$

が得られる。

これは漸近的に一様最強力な片側検定であり、ポアソン分布に基づく正確な一様最強力検定に対し、 n 大のときは十分良い近似を与える。

4.5.2 尤度比検定

帰無仮説 H_0 (4.56) を一般の複合対立仮説

$$H_2: \theta \neq \theta_0 \quad (\theta_0 \text{以外のすべての可能性を考える})$$

に対して検定するには

$$(4.59) \quad \frac{L(\tilde{\theta})}{L(\theta_0)} = \frac{f_{\tilde{\theta}}(\mathbf{y})}{f_{\theta_0}(\mathbf{y})} > c$$

のような棄却域が何らかの合理性を持つと考えられる。

なぜなら、もし θ_0 が真の値に等しければ $\tilde{\theta}$ が θ_0 の近傍にあることから $L(\tilde{\theta})$ と $L(\theta_0)$ は近い値をとることから、もし θ_0 が真の値とかけ離れていれば $L(\theta_0)$ が、 $L(\tilde{\theta})$ に対し、小さな値をとると考えられるからである(たとえば尤度関数の図 4.3 や 4.4 を参照)。

実際、棄却域(4.59)は最尤推定量に基づく両側検定と漸近的に同等であることが示される。この検定方式は(4.59)の形から尤度比検定 likelihood ratio test とよばれる。

以下、この検定を有意水準が α となるように定数 c を定めてみよう。いま、対数尤度 $\log L(\theta_0)$ を $\theta_0 = \tilde{\theta}$ のまわりでテイラー展開し、2次の項までとすればこれは

$$(4.60) \quad \begin{aligned} \log L(\theta_0) &= \log L(\tilde{\theta}) + \left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\tilde{\theta}} (\theta_0 - \tilde{\theta}) \\ &\quad + \frac{1}{2} \left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\tilde{\theta}} (\theta_0 - \tilde{\theta})^2, \\ &(\theta^* - \theta_0)(\theta^* - \tilde{\theta}) < 0 \end{aligned}$$

のようになる。この右辺第2項は最尤推定量の定義によって 0 である。一方、第3項は

$$\frac{1}{2} \cdot \frac{1}{n} \cdot \left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\tilde{\theta}} \cdot (\sqrt{n}(\tilde{\theta} - \theta_0))^2$$

と書き直すことができ、係数 $n^{-1} \cdot \partial^2 \log L(\theta)/\partial \theta^2|_{\theta=\tilde{\theta}}$ は H_0 の下で n が充分大のとき $-I_1(\theta_0)$ に確率収束するので

$$(4.61) \quad 2 \log \frac{L(\tilde{\theta})}{L(\theta_0)} \longrightarrow \{\sqrt{n}I_1(\theta_0)\}^2 (\tilde{\theta} - \theta_0)^2$$

がいえる。(4.61)の右辺は、 H_0 の下で、最尤推定量を θ_0 と漸近分散で規準化したものの二乗であるから漸近的に自由度 1 の分布 $\chi^2(1)$ に従う。これで尤度比検定の棄却域

$$(4.62) \quad 2 \log \frac{L(\tilde{\theta})}{L(\theta_0)} > \chi^2(1) \quad (= z_{\alpha/2}^2)$$

が導かれた。

なお、尤度比検定が最尤推定量 $\tilde{\theta}$ に基づく両側検定と漸近的に同等であることも示された。

例4.9 〈正規分布に対する平均の検定(続)〉 例4.7の設定で尤度比検定を求めよう。尤度は

$$L(\mu) = (2\pi\sigma_0^2)^{-\frac{n}{2}} \exp \left\{ -\frac{\sum(y_i - \mu)^2}{2\sigma_0^2} \right\}$$

で与えられ、 $\tilde{\mu} = \bar{y}$ である。これから

$$2 \log \frac{L(\bar{y})}{L(\mu_0)} = \frac{1}{\sigma_0^2} (\sum(y_i - \mu_0)^2 - \sum(y_i - \bar{y})^2) = \frac{n(\bar{y} - \mu_0)^2}{\sigma_0^2}$$

が得られる。この場合、尤度比検定は \bar{y} に基づく両側検定と正確に一致する。

例4.10 〈ポアソン分布に対する検定(続)〉 例4.8の設定で尤度比検定を求めよう。尤度は

$$L(\lambda) = e^{-n\lambda} \lambda^y / \prod_{i=1}^n y_i!$$

で与えられ、 $\tilde{\lambda} = \bar{y}$ である。これから

$$(4.63) \quad 2 \log \frac{L(\bar{y})}{L(\lambda_0)} = 2\{-n(\bar{y} - \lambda_0) + y \cdot \log(\bar{y}/\lambda_0)\} > z_{\alpha/2}^2$$

が得られる。 (4.63)において

$$\begin{aligned} (\bar{y}/\lambda_0) &= \log(1 + (\bar{y} - \lambda_0)/\lambda_0) \\ &= (\bar{y} - \lambda_0)/\lambda_0 - (1/2)\{(\bar{y} - \lambda_0)/\lambda_0\}^2 + \dots \end{aligned}$$

と展開し第3項以下を無視すると、 (4.63)は

$$\begin{aligned} &2\left(-n(\bar{y} - \lambda_0) + y(\bar{y} - \lambda_0)/\lambda_0 - \frac{y}{2}\{(\bar{y} - \lambda_0)/\lambda_0\}^2\right) \\ &= \frac{n(\bar{y} - \lambda_0)^2}{\lambda_0} \left(1 - \frac{\bar{y} - \lambda_0}{\lambda_0}\right) \end{aligned}$$

となり、 $(\bar{y} - \lambda_0)/\lambda_0$ が H_0 の近傍では $n^{-\frac{1}{2}}$ のオーダーで無視できることから、 (4.63) が最尤推定量 $\tilde{\lambda}$ に基づく両側検定の棄却域と漸近的に同等であることが確かめられる。

エフィシェント・スコア 帰無仮説 H_0 (4.56) の検定を構成するもう一つの考え方は、 尤度比の対数の母数による1階偏微分

$$\partial \log f_\theta(\mathbf{y}) / \partial \theta$$

に基づくものであって、 $\delta = \theta_1 - \theta_0 > 0$ でかつ十分小のとき、 尤度比の対数についての近似式

$$\log\{f_{\theta_1}(\mathbf{y})/f_{\theta_0}(\mathbf{y})\} \approx \partial \log f_\theta(\mathbf{y}) / \partial \theta|_{\theta=\theta_0} \cdot \delta$$

によっている。この $\partial \log f_\theta(\mathbf{y}) / \partial \theta$ をエフィシェント・スコア efficient score という。これに基づき

$$(4.64) \quad n^{-\frac{1}{2}} \cdot \frac{\partial \log f_\theta(\mathbf{y})}{\partial \theta} \Big|_{\theta=\theta_0} > c$$

のような棄却域を考えることができる。定数 c は有意水準の要請を満たすように定められる。

具体的に、エフィシェント・スコアに基づく有意水準 α の検定を作るには次のようにする。エフィシェント・スコアに $\theta = \theta_0$ を代入した式の H_0 の下での期待値は既に述べたように 0 となる。一方

$$n^{-\frac{1}{2}} \cdot \frac{\partial \log f_\theta(\mathbf{y})}{\partial \theta} \Big|_{\theta=\theta_0}$$

の漸近分散は $I(\theta_0)$ であるから、 (4.64) の左辺の H_0 の下での漸近分布は $N(0, I(\theta_0))$ に等しい。したがって、 $c = z_{\alpha/2} \sqrt{I(\theta_0)}$ とすればよいことがわかる。

それぞれ異なった考えから出発した、 4.5.1 および本項で述べた 2 つ(全部で 3 つ)の方法が漸近的には同等な検定方式を与えるのが興味深い。

例 4.11 〈正規分布の平均に対する検定(続)〉 例 4.7 の設定にエフィシェント・スコアの方法を適用すると例 4.7, 4.9 と正確に同じ結果が得られる。

例 4.12 〈ポアソン分布に対する検定(続)〉 例 4.8 の設定にエフィシェント・スコアの方法を適用する。尤度は例 4.10 に与えられている。それから

$$n^{-\frac{1}{2}} \frac{\partial \log L(\lambda)}{\partial \lambda} \Big|_{\lambda=\lambda_0} = n^{-\frac{1}{2}} \left(-n + \frac{y}{\lambda_0} \right) = \frac{\sqrt{n}(\bar{y} - \lambda_0)}{\lambda_0}$$

が得られる。 $I_1(\lambda_0) = \lambda_0^{-1}$ だから $\delta > 0$ に対する棄却域は

$$\sqrt{n}(\bar{y} - \lambda_0)/\lambda_0 > z_{\alpha/2} \sqrt{1/\lambda_0}$$

となる。これは例 4.8 の棄却域に一致する。

4.5.3 多次元母数の検定

例 4.2 の分割表を思い出して欲しい。 $a \times b$ 分割表の場合、本来の母数の次元は $ab - 1$ であるが、独立性の仮説

$$(4.65) \quad H_0: p_{ij} \equiv p_i \cdot p_j$$

の下での母数の個数は p_i のうち $a-1$ 個と p_j のうち $b-1$ 個の合計 $a+b-2$ である。このように本来 p 次元の母数 θ が $q (< p)$ 次元の母数 β によるある簡単な構造をしているとの帰無仮説

$$(4.66) \quad H_0: \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_p \end{pmatrix} = \begin{pmatrix} g_1(\beta) \\ \vdots \\ g_p(\beta) \end{pmatrix} \equiv g(\beta)$$

を検定したいことはよくある。ただし、 $\mathbf{g} = (g_1, \dots, g_p)$ は既知の関数である。とくに θ がある定数ベクトル θ_0 に等しいとする単純帰無仮説

$$(4.67) \quad H_0: \theta = \theta_0$$

は (4.66) で β が与えられた定数(したがって、未知母数の次元は 0)の場合として扱える。 H_0 (4.66) の検定にはよく用いられる 2 つの方法がある。

第一の方法は 4.5.2 節で述べた尤度比検定の多次元版であり、有意水準 α の棄却域が

$$(4.68) \quad 2 \log \{L(\tilde{\theta})/L(g(\tilde{\beta}))\} > \chi_{\alpha}^2(p-q)$$

で与えられる。ただし、 $\tilde{\theta}$ は仮説を課さない状態での θ の最尤推定量であり、 $\tilde{\beta}$ は H_0 (4.66) の下での最尤推定量である。読者はこの式を(4.59)および(4.62)のアノロジーで理解し、参照する χ^2 分布の自由度が $p-q$ になることに注意してほしい。単純帰無仮説(4.67)のときは $g(\tilde{\beta})$ を単に θ_0 とし、 $q=0$ と考えればよい。

例 4.13 〈二重分類の独立性〉 帰無仮説 H_0 (4.66)に対する尤度比検定方式を導き、表 4.3 のデータに適用する。

この場合の対数尤度は(4.10)に与えられているので、それを二通りの最尤推定量で評価すればよい。既に例 4.2 で求めたように、一般の場合の最尤推定量は $\tilde{p}_{ij}=y_{ij}/n$ 、帰無仮説 H_0 の下でのそれは $y_{i\cdot}y_{\cdot j}/n^2$ である。したがって

$$\begin{aligned} (4.69) \quad & 2 \log \{L(\tilde{\theta})/L(g(\tilde{\beta}))\} \\ & = 2 \{ \log L(\tilde{\theta}) - \log L(g(\tilde{\beta})) \} \\ & = 2 \left(\sum_i \sum_j y_{ij} \log \frac{y_{ij}}{n} - \sum_i \sum_j y_{ij} \log \frac{y_{i\cdot}y_{\cdot j}}{n^2} \right) \\ & = 2 \left(\sum_i \sum_j y_{ij} \log y_{ij} - \sum_i y_{i\cdot} \log y_{i\cdot} - \sum_j y_{\cdot j} \log y_{\cdot j} + n \log n \right) \end{aligned}$$

となる。 χ^2 分布の自由度は $p=ab-1$ 、 $q=a+b-2$ であることから

$$p-q=ab-1-(a+b-2)=(a-1)(b-1)$$

である。これで尤度比検定方式が得られた。

それでは表 4.3 に適用してみよう。(4.69)にデータを代入すると

$$\begin{aligned} & 2 \{(4 \log 4 + \dots + 6 \log 6) - (9 \log 9 + 18 \log 18 + 15 \log 15) \\ & \quad - (18 \log 18 + 9 \log 9 + 15 \log 15) + 42 \log 42\} = 0.19 \end{aligned}$$

となる。この値はきわめて小さく、自由度 $(3-1)(3-1)=4$ の χ^2 分布で全く有

表 4.5 期末試験成績のデータ(表 4.3 再掲)

代数 解析	優	良	可	計
優	4	2	3	9
良	8	4	6	18
可	6	3	6	15
計	18	9	15	42

本表は尤度比検定でも検定できる。

意とならない。この結果は、原データで代数(幾何)の成績によらず幾何(代数)の成績分布がほぼ一定という傾向が見られ、独立性を否定する証拠が何も示唆されないことをよく反映している。

第 2 の方法はエフィシェント・スコアに基づく方法の拡張である。この場合エフィシェント・スコアのベクトル

$$\frac{\partial \log L(\theta)}{\partial \theta} = \left(\frac{\partial \log L(\theta)}{\partial \theta_1}, \dots, \frac{\partial \log L(\theta)}{\partial \theta_p} \right)'$$

に基づく棄却域

$$(4.70) \quad n^{-1} \cdot \left(\frac{\partial \log L(\theta)}{\partial \theta} \right)' (I_1(\theta))^{-1} \left(\frac{\partial \log L(\theta)}{\partial \theta} \right)' \Big|_{\theta=\theta(\tilde{\beta})} > \chi_{\alpha}^2(p-q)$$

が得られるが、これも 1 次元パラメータの場合との類比で理解してほしい。この場合、(4.70)で見るようエフィシェント・スコアおよびフィッシャー情報量を H_0 の下での最尤推定量で評価すればよく、一般の最尤推定量 $\tilde{\theta}$ を求め必要のないことが特徴である。

この重要な特殊ケースとして、多項分布のセル確率に関し

$$H_0: p_i = p_i(\beta), \quad i=1, \dots, a$$

という仮説を想定する場合がある。たとえば、分割表の独立性の検定はその例である。この場合に(4.70)を変形するときわめて明解な形式

$$(4.71) \quad \chi^2 \equiv \frac{(y_1 - np_1(\tilde{\beta}))^2}{np_1(\tilde{\beta})} + \dots + \frac{(y_a - np_a(\tilde{\beta}))^2}{np_a(\tilde{\beta})}$$

が得られる。 np_i は第 i セルの期待度数であり、 $np_i(\tilde{\beta})$ はその仮説モデルの下でのあてはめ値(最尤推定量)である^{*}。したがって、(4.71)は実現値とあてはめ値の乖離を表す指標と解釈される。すなわちモデルの、実現値に対する適合度を測る指標とみなされるので、適合度 χ^2 統計量の名前でよばれる。これにはきわめて豊富な応用例があるので章をあらためて述べることにする。たとえば、(4.71)を分割表の独立性の検定に適用するときの具体的な形は(5.13)に、それを例 4.13 のデータに適用した結果は例 5.4 に与えられる。

*) 卷末広津(1982)

◆◆◆◆ 第4章 練習問題 ◆◆◆◆

4.1 † <心理学データ> 次のデータは標準音(*)より「強い」と感じられる、強さ(x)ごとの、相対度数のデータである。

強さ(デシベル)	同対数(x)	相対度数(p)
1.078	3.03	0.00
1.234	3.09	0.00
1.402	3.15	0.11
1.577	3.20	0.28
1.772*	3.25	0.48
1.972	3.30	0.71
2.169	3.34	0.83
2.375	3.37	0.91
2.579	3.41	0.95
2.793	3.45	0.98
3.011	3.48	1.00

(出典:スピアマン(1908) *Encyclopedia of statistics*. p. 817)

強さの対数 $\log x$ と相対度数 p のグラフを描け。

4.2 <幾何分布の最尤推定> 下記はあるコインを投げて表ができるまでの回数を調べた 100 回の実験の結果である。

回数	1	2	3	4	5	6	7	8	9	10	11	12	13以上
頻度	58	21	10	5	3	1	0	0	1	0	0	1	0

i) 表ができるまでの回数 x が幾何分布

$$f(x)=p(1-p)^{x-1}, \quad x=1, 2, \dots$$

に従うものとして、 p の最尤推定量が

$$\bar{p}=n/T$$

で与えられることを示せ。ここに n は実験の回数、 $T=\sum x_i$ はデータの合計である。このデータでの \bar{p} の値を求めよ。

ii) 幾何分布のフィッシャー情報量 $I_1(p)$ を求め、仮説 $p=p_0$ を対立仮説 $p>p_0$ に対して検定する有意水準 α の検定(4.58)を求め、このデータで $p_0=\frac{1}{2}$ を $\alpha=0.05$ で検定せよ。

iii) ii)の検定では \bar{p} が大きいとき、すなわち T が小さいとき仮説が棄却される。 T の分布を平均と分散の等しい正規分布 $N(n/p, n(1-p)/p^2)$ で近似することにより、近似的な棄却域を導け。このデータで $p_0=\frac{1}{2}$, $\alpha=0.05$ のときの結論を述べよ。

iv) 仮説 $p=p_0$ を $p\neq p_0$ に対して検定する尤度比検定の棄却域(4.62)を求めよ。このデータで $p_0=\frac{1}{2}$ を $\alpha=0.05$ で検定せよ。

v) ii)から iv)の結果を比較して考察せよ。

4.3 <標準偏差の最尤推定> x_1, \dots, x_n が平均 0, 分散 σ^2 の正規分布からの観測値であるとき

i) σ の最尤推定量 $\hat{\sigma}$ を求めよ。

ii) n が十分大きいときの $\hat{\sigma}$ の分散を求めよ。

4.4 <指数分布の最尤推定> 次のデータはある電子部品 $n=10$ 個を通常より過酷な条件で試験したときの故障時間 [min] である。

523	971	878	130	696
201	1643	491	665	431

(架空例)

i) 故障時間 T が指数分布

$$f(t)=\lambda e^{-\lambda t} \quad (t \geq 0)$$

に従うものとして、 λ の最尤推定量 $\tilde{\lambda}$ の値を求めよ。

ii) 一般に母数 θ の最尤推定量を $\tilde{\theta}$ とするとき、 θ の関数 $g(\theta)$ の最尤推定量は $g(\tilde{\theta})$ で与えられる。このことを用いて、故障時間が $t_0=1000$ [min] を超える確率の最尤推定値を求めよ。

4.5 <一様分布の最尤推定> 区間 $[0, \theta]$ 上の一様分布からの観測値 x_1, \dots, x_n による θ の最尤推定量を求めよ。

第5章 適合度検定

データがある分布、たとえばポアソン分布に適合しているかどうか、あるいはデータに簡単な統計モデルが適合するか否かはそのデータの発生メカニズムを知り、適切な推論を実行する上で大切なことである。この章ではそのための適合度検定の方法とその適用例を中心に説明しよう。この方法は自然科学、人文科学、社会科学など、幅広く用いられているので、興味・関心に応じて活用することができる。

5.1 χ^2 適合度検定

多項分布の各項の生起確率 p_1, p_2, \dots, p_a がある一定の法則に従っているという仮説の検定は一般に適合度検定 test of goodness of fit とよばれており、いろいろな応用がある。

その検定法としては、4.5節で述べた尤度比検定(4.68)やエフィシェント・スコアに基づく検定(4.71)が用いられる。とくに後者は検定統計量の形から χ^2 適合度検定 χ^2 test of goodness-of-fit の名で頻用されている。(4.71)において $np_i(\tilde{\beta})(i=1, 2, \dots, a)$ は仮説モデルの下でのあてはめ値であり、とくに帰無仮説が単純仮説の場合などでは理論度数 theoretical frequency というより名がふさわしい。そのため

$$(5.1) \quad \chi^2 = \sum \frac{(\text{観測度数} - \text{理論度数})^2}{\text{理論度数}}$$

と表すことも多い。いずれにせよ、観測値へのモデルのあてはまり度を測ると

表5.1 0, 1, ..., 9 の乱数発生度数

乱 数	0	1	2	3	4	5	6	7	8	9	計
度 数	10	9	7	7	9	6	10	17	14	11	100
正	正	正	正	正	正	正	正	正	正	正	
正	正	丁	丁	正	一	正	正	正	正	正	
正	正	丁	丁	正	正	正	正	正	正	正	
正	正	正	正	正	正	正	正	正	正	正	
丁	丁	丁	丁	丁	丁	丁	丁	丁	丁	丁	

一様乱数は本当に一様(ランダム)か、これを検証するために 100 個発生させ、各数の出た度数を記したものである。一様に出現すると期待される結果の一様性の統計的検証の例は、適合度検定の典型的な場合であって、乱数の発生のよりよいメカニズムは第 11 章を参照。

なお、度数を数える‘正’は外国では用いせず、アメリカなどでは、冊が用いられている。

いう意味で適合度統計量 goodness-of-fit statistic というより方がふさわしい。すでに述べたように χ^2 の帰無仮説の下での漸近分布は χ^2 分布で、その自由度は

$$(5.2) \quad \nu = (\text{多項分布の自由度}) - (\text{あてはめたモデルの母数の個数}) \\ = (\text{項の数 } (a) - 1) - (\text{あてはめたモデルの母数の個数})$$

で与えられる。

一方、尤度比検定も有力な方法であり、帰無仮説(あてはめるモデル)の近傍では適合度 χ^2 と漸近的に同等な検定を与える。

例 5.1 <セル確率の一様性検定> 表 5.1 は 0, 1, ..., 9 の一様乱数を 100 個生成した結果である。乱数が正しく作られているかどうかを χ^2 適合度検定によって調べてみよう。

この場合、一様乱数が正しく発生されていれば

$$n=100, \quad p_0=p_1=\cdots=p_9=1/10$$

の多項分布に従うはずである。そこで多項分布の各セルの生起確率に関する一様性の帰無仮説

$$(5.3) \quad H_0 = p_0 = p_1 = \cdots = p_{10} = 1/10$$

を χ^2 適合度統計量によって検定することができる。

それには(5.1)を計算すればよいが、理論度数はすべてのセルについて

$$np_k = 100 \cdot \frac{1}{10} = 10$$

である。したがって、(5.1)は

※ 999999999 でも 10 億桁ではランダム※

円周率[†] π を乱数として用いようとする考えはよく試みられている。もちろん、その前提として、

$$\pi = 3.1415926535\cdots$$

が十分にランダムにならんだ数字であることが確認されねばならない*。

ところで、学生に‘ランダムに数字を 100 個書け’というと、誰でも何とかして数字をばらつかせようという意識が働いて、222 という重複数字や 567 という連続数字をことさらに避けようとするものであるが、しかし本来ランダムであれば、特定の選び方が意識的に避けられてはならないはずである。

実際、日本で最近 π が 10 億桁計算されたという実績があるが、これらの数字がランダムであることが適合度検定をはじめとして総合的方法で確かめられている。にもかかわらず、驚いたことに、5 億 6466 万 5206 桁目(下にその付近を示す)から‘9’が 9 個(nine nines)も続き

$$999999999$$

のように並ぶ。さらに、3 億 8698 万 0412 桁目からは‘6’が実際に 10 個続く。

奥深い所にはこんな神秘がある。それだけではなく、最初の 1000 桁の中できさえすでに 762 桁目から 6 個の‘9’(999999)が続いている。このように π はランダムには見えないのに統計学的にはランダムであるという不思議な性質を持っている(出典: 金田康正『 π のはなし』東京図書、1991)。

21858647503620761322503876977908585280703141442546
10428565190082957740713494770317918594586555725861
0478263748336914810457164982677746092685194551128
89367999999999419275842728155297933735757536788421
16450051116954672987792452984775309206278275092967
80631138802369294247674936518450125907807439823059
74963808221887341446597347815792806230194792117217

*) 本シリーズ第 1 卷『統計学入門』第 12 章練習問題 12.5、および本章練習問題 5.7 参照。

$$\chi^2 = \frac{(10-10)^2}{10} + \frac{(9-10)^2}{10} + \cdots + \frac{(11-10)^2}{10} = \frac{102}{10} = 10.2$$

となる。いま、あてはめたモデルの母数(パラメータ)の数は 0 だから、 χ^2 分布の自由度は(5.2)より

$$\nu = 10 - 1 = 9$$

と求まる。自由度9の χ^2 分布 $\chi^2(9)$ の上側25%点は $\chi^2_{0.25}(9)=11.39$ 、50%点は $\chi^2_{0.5}(9)=8.34$ で、実測値 $\chi^2=10.2$ はちょうどその中間にあつた。この結果、通常の有意水準で、セル確率の一様性は棄却されない。ただし、セル確率の一様性は一様乱数の一つの側面でしかない。正しい一様乱数であるためにはどのくらい続く2つをとっても対として一様に発生している等々のことが必要であることを注意しておこう。

いま、多項分布の尤度

$$(5.4) \quad L(\mathbf{p}) = \frac{n!}{y_1! \cdots y_a!} p_1^{y_1} \cdots p_a^{y_a}$$

に基いてあらためて尤度比検定(4.68)を作るには次のように考える。まず、特別な構造を仮定しないときには最尤推定量

$$\tilde{p}_k = y_k/n, \quad k=1, \dots, a$$

が得られる。一方、セル確率が

$$H_0: p_k = p_k(\beta), \quad k=1, \dots, a$$

という構造であるとの帰無仮説の下では、母数 β の最尤推定量を $\tilde{\beta}$ として $p_i(\tilde{\beta})$ という最尤推定量が得られる。これらを(5.3)に代入して(4.68)を作る多項係数は相殺されて

$$(5.5) \quad 2 \log \left[\left\{ \left(\frac{y_1}{n} \right)^{y_1} \cdots \left(\frac{y_a}{n} \right)^{y_a} \right\} / \{ (p_1(\tilde{\beta}))^{y_1} \cdots (p_a(\tilde{\beta}))^{y_a} \} \right] \\ = 2 \sum_{k=1}^a y_k \log \frac{y_k}{np_k(\tilde{\beta})}$$

という公式が得られる。これを分割表(二重分類)の独立性の検定に適用したのが公式(4.69)にほかならない。(5.5)を尤度比

$$\lambda = \frac{L(p_1(\tilde{\beta}), \dots, p_a(\tilde{\beta}))}{L(y_1/n, \dots, y_a/n)}$$

を用いて $-2 \log \lambda$ と表す習慣があり、ここでもそれに習うこととする。 $-2 \log \lambda$ を(5.1)に対応して表現すると

$$(5.6) \quad -2 \log \lambda = 2 \sum (\text{観測度数}) \log \left(\frac{\text{観測度数}}{\text{理論度数}} \right)$$

のようになる。 $-2 \log \lambda$ はモデルが真であるとの帰無仮説の下でやはり漸近

的に χ^2 分布に従い、その自由度は(5.2)で与えられる。

例5.2 (続) 表5.1のデータに関し、セル確率の一様性の帰無仮説 H_0 (5.3)の尤度比検定を行おう。今度は(5.6)を用いて次のように統計量を計算する。

$$\begin{aligned} -2 \log \lambda &= 2 \left(10 \log \frac{10}{10} + 9 \log \frac{9}{10} + \cdots + 11 \log \frac{11}{10} \right) \\ &= 2(10 \log 10 + 9 \log 9 + \cdots + 11 \log 11 - 100 \log 10) \\ &= 2(235.0833 - 230.2585) \\ &= 2 \cdot 4.8248 = 9.65 \end{aligned}$$

この値は先の χ^2 統計量とよく一致しており、もちろん結論は変わらない。以下の各節で適合度検定の応用例をいくつか示す。

5.1.1 分布の適合度検定

表5.2はある原子核崩壊で一定時間(7.5秒)に対極に到達した放射性物質からの α 粒子を2608回観測したデータで、ポアソン分布がよく適合する例とし

表5.2¹ 到達 α 粒子データへのポアソン分布あてはめ
(ラザフォード、ガイガー)

到達数 k	観測度数 y_k	理論度数 \hat{y}_k
0	57	54.399
1	203	210.523
2	383	407.361
3	525	525.496
4	532	508.418
5	408	393.515
6	273	523.817
7	139	140.325
8	45	67.882
9	27	29.189
10以上	16	17.075
計	2608	2608.000

(原出典: ラザフォード、ガイガー)

測定は毎7.5秒の到達カウント数。Po(3.87)にあてはめた理論度数をかけた。ポアソン分布(ポアソンの小数の法則)へのあてはまりは從来ボルトキーヴィッヂの例(第1章参照)によって説明されるが、このデータは、純粹に自然現象からのものであり、ついでよく引用される。

て有名なものである。

ポアソン分布は、第1章で述べたように、時間的あるいは空間的に生起確率が一定のランダム事象にあてはまる分布であるから、原子核崩壊のように時間的に状態の変化する事象には本来適合しないはずと思われるかも知れない。しかししながら半減期に数年を要するものでは、この例程度の観測時間内では状態一定と見なし得る。実際にポアソン分布のあてはまりを見てみよう。

例 5.3 〈ポアソン・モデルの適合度検定〉 表5.2の放射性物質データに対し、ポアソン・モデルの適合度検定を行ってみよう。表5.2にあるように $k \geq 10$ をまとめて $n=2608$ の11項分布が観測されたと考える。この各項の生起確率 p がポアソン分布に従うという仮説は

$$(5.7) \quad H_0: p_k = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k=0, 1, \dots, 9; \quad p_{10} = 1 - p_0 - p_1 - \dots - p_9$$

という多項分布 multinomial distribution のモデルを想定することと同値である。この仮説はちょうど先の例5.1で一様性の帰無仮説(5.3)を想定したことに対応する。ただし未知母数 λ が含まれている点が異なっている。母数 λ の推定量は理論上(5.7)を、多項分布としての尤度

$$(5.8) \quad 2608! \prod_{k=0}^{10} p_k^{y_k} / y_k!$$

に代入して、その最尤推定量として求めるべきであるが、ここでは便宜上ポアソン分布としての有効推定量

$$(5.9) \quad \tilde{\lambda} = \sum_{k=0}^{\infty} k y_k / n = (10086 / 2608) \div 3.87$$

を用いる。両者の違いは $k \geq 10$ をまとめるか否かの相違であり、この例の場合その影響は小さい。 $k \geq 10$ をまとめる必要があるのは適合度統計量の漸近特性のためである(漸近分布が χ^2 分布でよく近似されるためには各セルの期待度数が3以上であることが望ましいといわれている)。

さて、(5.7)に $\tilde{\lambda}=3.87$ を代入したものを $\bar{p}_k (k=0, \dots, 10)$ と置き、それから $n \bar{p}_k$ を計算すると表5.2右端のような理論(あてはめ)度数が得られる。これから(5.1)が次のように計算される。

$$\begin{aligned} \chi^2 &= \frac{(57-54.399)^2}{54.399} + \dots + \frac{(16-17.075)^2}{17.075} \\ &= 12.885 \end{aligned}$$

自由度はあてはめたモデルの母数が λ の1個だけだから(5.2)より

$$\nu = 11 - 1 - 1 = 9$$

である。 $\chi^2=12.885$ は自由度9の χ^2 分布 $\chi^2(9)$ として上側確率約0.17の点にあたり、ポアソン・モデルの適合は、通常の有意水準では棄却されない。すなわち、仮説は有意ではない。実際、表5.2では $k=8$ を除き、 y_k と \hat{y}_k が極めてよく一致しているのが見てとれる。

5.1.2 分割表における独立性検定

表5.3はクラス42名を代数と解析の期末試験成績によって分類したものである(表4.3再掲)。この例のように各個体を2種の属性に従って2重分類して得られる表を2次元 two-way の分割表 contingency table という。

表5.3をよく見ると、例4.13で見たように、代数で「優」を取った9人の解析の成績分布と、「良」を取った18人の解析の成績分布は完全に比例し、「可」の15人についてもほぼ同様の傾向が見られる。すなわち常識的に期待される両科目間の正相関はありそうもない。このような2重分類の独立性 independence を統計的に検証するには生起確率に関する独立モデル

$$(5.10) \quad H_0: p_{ij} = p_{i \cdot} \cdot p_{\cdot j}$$

の適合度検定を行えばよい。これを独立性の検定 test of independence という。この場合の尤度比検定は例4.13すでに求めた。ここでは χ^2 適合度検定を導こう。一般にX分割表の尤度は(4.5)に与えられている。特に帰無仮説

表5.3' 2次元分割表：期末試験成績

分析 代数	優	良	可	計
優	4(3.86)	2(1.93)	3(3.21)	9(9.00)
良	8(7.71)	4(3.86)	6(6.43)	18(18.00)
可	6(6.43)	3(3.21)	6(5.36)	15(15.00)
計	18(18.00)	9(9.00)	15(15.00)	42

(出典：広津)

2次元分割表の典型的な例である。かっこ内は独立モデルの理論度数を示した。

(5.10) のもとで最尤推定量は、例 4.2 に見るように

$$(5.11) \quad \tilde{p}_{ij} = \frac{y_{i\cdot}}{n} \cdot \frac{y_{\cdot j}}{n}$$

で与えられる。これからセル (i, j) のあてはめ度数(仮説の下での理論度数)

$$(5.12) \quad \hat{y}_{ij} = n \cdot \frac{y_{i\cdot}}{n} \cdot \frac{y_{\cdot j}}{n} = \frac{y_{i\cdot} y_{\cdot j}}{n}$$

が得られ、 χ^2 適合度統計量

$$(5.13) \quad \chi^2 = \sum_{i=1}^a \sum_{j=1}^b \left\{ \left(y_{ij} - \frac{y_{i\cdot} y_{\cdot j}}{n} \right)^2 \Big/ \left(\frac{y_{i\cdot} y_{\cdot j}}{n} \right) \right\}$$

が導かれる。自由度は $p.. \equiv 1$ の制約から H_0 のもとでの母数が $(a-1)+(b-1)$ だけあるので

$$(5.14) \quad \nu = ab - 1 - (a-1+b-1) = (a-1)(b-1)$$

となる。

例 5.4 <独立性の検定> 表 5.3 の試験成績データについて独立性の検定を行ってみよう。 (5.12) によるあてはめ度数は表 5.3 の括弧内に示してある。一見して独立モデルはよくあてはまっている。実際、適合度統計量は

$$\chi^2 = \frac{(4-3.86)^2}{3.86} + \cdots + \frac{(6-5.36)^2}{5.36} = 0.19$$

となり、有意性は問題にならない。この結果は例 4.13 の尤度比検定の結果

$$-2 \log \lambda = 0.19$$

と丸めの誤差の範囲で一致している。

5.2 2つの二項分布の比較

表 5.4 はある工程の改善前後の不良率を比較するためにそれぞれ標本を取り、良・不良に従って分類したものであり、生産工程でよく見られる例である。これと類似の例は他にもいろいろな分野に見られ、たとえば表 5.5 のようなデータが医学分野でよく見られる。これらでは処置の前後で発生比率を比較し、処置の効果があったか否かを判定することが目的になる。

すなわち、表 5.4 については工程の改善前後の不良率をそれぞれ p_1, p_2 、ま

表 5.4 ドラム缶製造工程の溶接不良の比較

工程 品種	良 不 良	
	改善前	改善後
改善前	2,422	439
改善後	2,892	447

(出典:森口)

改善前後の比較であるから、2行を比べることになり、表 5.6 の場合に帰着する。

表 5.5 予防注射の効果

注射 有無	かからない かかった	
	したた	しない
かからない	1,625	5
かかった	1,022	11

(出典:グリーンウッド, ュール)

これも表 5.4 のように、ある効果の'前後'比較を横比で見るものである。表 5.6 に帰着する。結論は例 5.6 で与えられている。

表 5.6 2つの二項分布の比較

群	2 値応答		群	生起確率(p_{ij})	
	0	1		0	1
1	y_{11}	y_{12}	1	$1-p_1$	p_1
2	y_{21}	y_{22}	2	$1-p_2$	p_2

上の表 5.4, 表 5.5 とも、一般的な形のこの表に帰着する。

た表 5.5 については注射の有無による有効率(または無効率)を p_1, p_2 として帰無仮説

$$H_0: p_1 = p_2$$

の検定に興味が持たれる。

これはたとえば次のような考え方で検定することができる。いま、表 5.4, 表 5.5 のようなデータを一般に表 5.6 のように表してから、生起確率の推定量を

$$\tilde{p}_1 = y_{12}/n_1, \quad \tilde{p}_2 = y_{22}/n_2$$

とすると、 $\tilde{p}_1 - \tilde{p}_2$ は帰無仮説

$$(5.15) \quad H_0: p_1 = p_2 (= p)$$

の下で期待値は 0, 分散は

$$V(\tilde{p}_1 - \tilde{p}_2) = V(\tilde{p}_1) + V(\tilde{p}_2)$$

$$= \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \cdot p(1-p)$$

となり、また漸近分布は中心極限定理より、正規分布

$$N\left(0, \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \cdot p(1-p)\right)$$

となる。いま帰無仮説 H_0 のもとで考えるとして、共通の確率 p を

$$(5.16) \quad \tilde{p} = \frac{y_{12} + y_{22}}{n_1 + n_2}$$

で推定しよう。たとえば、予防注射のデータでは

$$\tilde{p}_1 = \frac{5}{1625+5} = 0.003, \quad \tilde{p}_2 = \frac{11}{1022+11} = 0.011, \quad \tilde{p} = \frac{16}{1630+1033} = 0.006$$

などとなる。

標準化した統計量

$$u = \frac{\tilde{p}_1 - \tilde{p}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \cdot \tilde{p}(1-\tilde{p})}}$$

を定義すると、 u は H_0 の下で漸近的に標準正規分布 $N(0, 1)$ に従う。したがって u^2 の漸近分布は自由度 1 の χ^2 分布 $\chi^2(1)$ となる。ところで u^2 は簡単な式変形により

$$(5.17) \quad u^2 = \frac{y_{..} \cdot (y_{11}y_{22} - y_{12}y_{21})^2}{y_{1..}y_{2..}y_{.1}y_{.2}}$$

と表される。ところがこれは(5.13)において $a=b=2$ と置いたものにほかならない(例 5.5 参照)。これはこの節の設定での確率分布(尤度)は二項分布の積

$$(5.18) \quad L(p) = \prod_{i=1}^2 \left\{ n_i! \prod_{j=1}^2 (p_{ij})^{y_{ij}} / y_{ij}! \right\}$$

のように表すことができ、分布の核となる部分は(4.5)と同じになることからきている。さらに独立性の仮説を

$$(5.19) \quad H_0': \quad \frac{p_{ij}}{p_{i..}} \equiv p_{.j} \quad (i \text{ によらない})$$

のよう表すとそれが本質的に $H_0: p_1 = p_2$ と同値であることも明らかである。以上から二項分布の同等性検定にも、分割表の独立性検定の χ^2 統計量がそのまま使えることがわかった。

例 5.5 <2×2 の場合> (5.13)において $a=b=2$ としたものが(5.17)と一致することを確かめてみよう。たとえば $i=j=1$ の項は

$$\begin{aligned} & \left(y_{11} - \frac{y_{1..}y_{.1}}{n} \right)^2 / \frac{y_{1..}y_{.1}}{n} \\ & = \frac{(y_{11}(y_{11} + y_{12} + y_{21} + y_{22}) - (y_{11} + y_{12})(y_{11} + y_{21}))^2}{ny_{1..}y_{.1}} \\ & = \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{ny_{1..}y_{.1}} \end{aligned}$$

のよう変形される。その他の i, j についても同様に考えて

$$\left(y_{ij} - \frac{y_{i..}y_{.j}}{n} \right)^2 / \frac{y_{i..}y_{.j}}{n} = \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{ny_{i..}y_{.j}}$$

が得られる。これらを(5.13)に代入すると

$$\begin{aligned} \chi^2 &= \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{n} \left(\frac{1}{y_{1..}y_{.1}} + \frac{1}{y_{1..}y_{.2}} + \frac{1}{y_{2..}y_{.1}} + \frac{1}{y_{2..}y_{.2}} \right) \\ &= \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{n} \cdot \frac{y_{2..}y_{.2} + y_{2..}y_{.1} + y_{1..}y_{.2} + y_{1..}y_{.1}}{y_{1..}y_{.2}y_{.1}y_{.2}} \\ &= \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{n} \cdot \frac{y_{2..} \cdot n + y_{1..} \cdot n}{y_{1..}y_{.2}y_{.1}y_{.2}} \\ &= \frac{(y_{11}y_{22} - y_{12}y_{21})^2}{n} \cdot \frac{n^2}{y_{1..}y_{.2}y_{.1}y_{.2}} \\ &= \frac{n(y_{11}y_{22} - y_{12}y_{21})^2}{y_{1..}y_{.2}y_{.1}y_{.2}} \end{aligned}$$

つまり、(5.17)が得られる。この結果はよく知られている。

例 5.6 <二項分布の比較> 表 5.5 の予防注射データについてその有効性に関する帰無仮説を検定してみよう: 公式(5.14)により

$$\chi^2 = \frac{2663(1625 \cdot 11 - 1022 \cdot 5)^2}{1630 \cdot 1033 \cdot 2647 \cdot 16} = 6.08$$

が得られる。自由度は $\nu = (2-1)(2-1) = 1$ であり、 $\chi^2_{0.05}(1) = 3.84$ であるから、実現値 6.08 に基づいて帰無仮説は有意水準 5% で棄却される(5% 有意である)。すなわち、予防注射の効果はあったと考えられる。

なお、このデータは χ^2 が当初自由度 3 の χ^2 分布に従うとして検定し、 $\chi^2_{0.05}(3) = 7.82$ であることから有意差無しと結論づけられていた^{*)}のを、フィッシ

^{*)} グリーンウッドおよびユール

ヤーが自由度計算の誤りを指摘し、有意差有りと訂正したものである。これ以前に分割表の自由度を $\nu = ab - 1$ とする説と $\nu = (a-1)(b-1)$ とする説の間に 20 年に及ぶ論争があった。

5.3 多項分布の一様性検定

5.3.1 一様性検定

確率が等しいことの検定は表 5.7 のような 3つ以上の多項分布の一様性の検定 test of homogeneity にもそのまま拡張される。それは数学的には多項分布

表 5.7 職業別に分類した初登院時のガンの病状(ガンの重症度データ)

職業	病状	軽症	中症	重症	計
1. 専門的、技術的職業 (土木技術者、教員、医師等)		148(123.3) 148.5	444(473.9) 452.4	86(80.8) 77.1	678
2. 管理職		111(93.1) 112.1	352(357.9) 341.6	49(60.0) 58.2	512
3. 事務専従者 (会計事務員、タイピスト等)		645(524.6) 631.4	1,911(2,015.7) 1,924.4	328(343.7) 328.1	2,884
4. 販売従事者		165(191.9) 160.7	771(737.4) 764.0	119(125.7) 130.3	1,055
5. 農林・漁業、採鉱従事者		383(458.9) 384.3	1,829(1,763.4) 1,827.2	311(300.6) 311.5	2,523
6. 運輸、通信従事者		96(79.3) 95.5	293(304.7) 290.9	47(52.0) 49.6	436
7. 技能工 (製鉄工、自動車修理工等)		98(88.4) 106.4	330(339.7) 324.3	58(57.9) 55.3	486
8. 生産工程従事者、単純労働者		199(223.4) 187.1	874(858.3) 889.3	155(146.3) 151.6	1,228
9. サービス業		59(52.4) 63.1	199(201.3) 192.2	30(34.3) 32.8	288
10. 無職		262(330.7) 276.9	1,320(1,270.7) 1,316.6	236(216.6) 224.5	1,818
計		2,166	8,323	1,419	11,908

(出典: 広津(1983))

()内上段は独立モデルのあてはめ、下段は次項で述べる自由度 1 交互作用モデルのあてはめ。
「ガンの早期発見」に大きなかかわりのある興味深いデータである。各職業ごとの初期登院時の病状(軽症、中症、重症)に差があれば、職業ごとの早期発見システムに何らかの差があると考えられるからである。結論からいうと、重症度分布には差があると考えられる(表 5.8, 5.9 に詳しい分析がある)。

の一様性 homogeneity の帰無仮説

$$H_0: (p_{i1}, p_{i2}, p_{i3}) = (p_1, p_2, p_3), \quad i=1, \dots, 10$$

の検定として定式化される。これは 5.2 節の 0, 1 応答をより一般化したものと考えられ、同じ論理により検定統計量として(5.1)の適合度の χ^2 統計量が導かれる。

例 5.7 <一様性の適合度検定> 表 5.7 は各企業の検診で発見された後、国立ガンセンターに送り込まれたガン患者を企業ごとに初登院時の重症度で分類して示したものである。もし企業ごとにガンの早期発見システムに差があればそれは初登院時の重症度分布の差として現れるはずである。そこでそのような差がないとする帰無仮説の検定に興味を持たれる。

このデータに関して χ^2 適合度検定を行ってみよう。 χ^2 の値は

$$\begin{aligned} \chi^2 &= \frac{(148-123.3)^2}{123.3} + \frac{(444-473.9)^2}{473.9} + \dots + \frac{(236-216.6)^2}{216.6} \\ &= 96.39 \end{aligned}$$

となる。自由度は $a=10$, $b=3$ から $(10-1)(3-1)=18$, $\chi^2_{0.005}(18)=37.16$ であるから実現値 96.39 は極めて高度に有意である。すなわち企業により重症度分布には差があると考えられる。ただし、一様性検定によって高度の不均一性が示されてもデータ解析としては不十分である。一步進めて、どのような不均一性があるのか、あるいは独立モデルに替ってどのようなモデルが適合するのかを明らかにしなければならない場合のために、その手法について説明しよう。

5.3.2 χ^2 統計量の分解*

(5.10) の総括的な χ^2 統計量をある 2 行間の差を表す成分、あるいは行をいくつかの群に分類したとき群間の差を表す成分に分解することができる。たとえば一般性を失うことなく、 $\{1, 2, 3, \dots, p+q\}$ を分けた 2 群を

$$I_1 = \{1, \dots, p\}, \quad I_2 = \{p+1, \dots, p+q\}$$

とするとき、群間の差を表す χ^2 成分は

*) やや進んだ内容であるが、手法がわかればよい。簡単な場合は自分で確かめることも可能である。

$$(5.20) \quad \chi^2(I_1; I_2) = \frac{y_{..} \sum_{j=1}^b (\sum_{l'}^{(1)} y_{lj}/\sum_{l'}^{(1)} y_{l..} - \sum_{l'}^{(2)} y_{lj}/\sum_{l'}^{(2)} y_{l..})^2 / y_{..}}{(1/\sum_{l'}^{(1)} y_{l..}) + (1/\sum_{l'}^{(2)} y_{l..})}$$

で与えられる。ここで、 l に関する和は、 I_1, I_2 の上でそれぞれ

$$\sum_{l'}^{(1)} \equiv \sum_{l=1}^p, \quad \sum_{l'}^{(2)} \equiv \sum_{l=p+1}^q$$

のようにとられるものとする。これを用いて全体の χ^2 をたとえば

$$(5.21) \quad \chi^2 = \chi^2(1; 2) + \chi^2(1, 2; 3) + \cdots + \chi^2(1, \dots, a-1; a)$$

のよう分解することができる。行を入れ替えることにより、 $a=3$ の場合に

$$(5.22) \quad \begin{aligned} \chi^2 &= \chi^2(1; 2) + \chi^2(1, 2; 3) = \chi^2(1; 3) + \chi^2(1, 3; 2) \\ &= \chi^2(2; 3) + \chi^2(2, 3; 1) \end{aligned}$$

が成り立つことがわかるし、また $a=4$ では

$$(5.23) \quad \chi^2 = \chi^2(1; 2) + \chi^2(3; 4) + \chi^2(1, 2; 3, 4)$$

のよう分解することもできる(読者は $a=3$ のとき(5.22)を確かめてほしい)。

そこであらゆる組合せについてこのような χ^2 成分を計算し、群内 within groups は均一に、かつ群間 between groups の χ^2 成分はなるべく大きくなるように群分けすることが考えられる。ただし、かかる群分けの有意性を統計的に評価するには少し注意が必要である。

H_0 の下で個々の χ^2 成分は、漸近的に自由度 $\nu=b-1$ の χ^2 分布 $\chi^2(b-1)$ に従う。その最大値を探した場合には統計量としてもそれら χ^2 成分の最大統計量として評価しなければならない。そうしないと、いわゆる‘後知恵解析’となって多大の偽陽性(言い過ぎ)を生じてしまう。この場合、最大統計量はウィッシャート行列(説明は略す)の最大固有値の分布に従うことがわかっており、また、探索的に群分けを行うアルゴリズムも提案されているが、やや高度な議論となるので詳細は略する。

ここでは結果だけの紹介にとどめるとし、表 5.7 のガンの重症度のデータについて各 2 行ごとに、(5.20) を適用した結果を表 5.8 に示す。ただし χ^2 値の小さい組合せは近く、大きいものは遠くなるように並べ直してある。以下は、先へ進むと表 5.7 からこのような分析も可能という興味深い事例である。

さて、表 5.8 は明らかに職業に関する 2 群

$$I_1 = \{10, 5, 4, 8\}, \quad I_2 = \{7, 9, 2, 1, 6, 3\}$$

表 5.8 表 5.7 のデータに対する行間の χ^2 値: 職業の群分け

職業	10	5	4	8	7	9	2	1	6	3
10	0	0.73	2.20	1.59	8.55	6.85	16.4	18.6	14.1	47.6
5		0	0.81	0.74	6.75	5.22	13.5	16.8	11.8	46.7
4			0	1.25	5.14	3.58	8.68	12.8	8.49	24.4
8				0	3.67	3.45	9.08	9.62	7.62	22.1
7					0	0.40	1.51	0.80	0.71	1.38
9						0	0.26	1.45	0.33	1.01
2							0	2.82	0.38	1.68
1								0	0.92	0.93
6									0	0.18
3										0

(出典: 広津(1983))

{10, 5, 4, 8} : 無職 農林・漁・採鉱從事者、販売從事者、生産工程從事者・単純労働者

{7, 9, 2, 1, 6, 3} : 技能工、サービス業、事務専從者、専門的・技術的職業、運輸・通信從事者、管理職

への群分けを示唆する。実際この 2 群間の χ^2 は

$$\chi^2(10, 5, 4, 8; 7, 9, 2, 1, 6, 3) = 87.74$$

となり、ただ一つの成分で全体の $\chi^2 = 96.39$ の 91.03% を説明している*。

一方、列に関し同様の考え方を適用した結果からは

$$J_1 = \{1\}, \quad J_2 = \{2, 3\}$$

という群分けが得られた。これらの結果は次のような自由度 1 の交互作用モデルを示唆する。それは、 I_1, I_2 および J_1, J_2 の各組合せを区別する新しい母数

$$\lambda_{I_1 J_1}, \lambda_{I_1 J_2}, \lambda_{I_2 J_1}, \lambda_{I_2 J_2}$$

を導入して、完全独立モデル $p_{ij} = p_{i..} p_{..j}$ を

$$(5.24) \quad p_{ij} = p_{i..} p_{..j} \lambda_{ij}, \quad i \in I, j \in J$$

と修正した交互作用のモデルである。ただし、 I は I_1, I_2 の 2 通り、 J は J_1, J_2 の 2 通りをとるので、 λ_{ij} のうち独立な母数は $(2-1) \cdot (2-1) = 1$ 個であることに注意せよ。

モデル(5.24)を仮定したときのセルの度数の最尤推定量は

$$\tilde{m}_{ij} = y_{i..} y_{..j} Y_{ij} / (R_i C_j), \quad i \in I, j \in J$$

で与えられる。ただし

*) データの解析は卷末文献による。

$$Y_{IJ} = \sum_{i \in I} \sum_{j \in J} y_{ij}, \quad R_I = \sum_{i \in I} y_{i..}, \quad C_J = \sum_{j \in J} y_{.j}$$

である。

表5.7のガンの重症度のデータにモデル(5.24)をあてはめた結果は表5.7の()内下段に示してある。独立モデルにただ一つの母数を追加しただけで現データと極めてよい一致が得られる。たとえば、(1,1)セルに対しては

$$y_{11} = 678, \quad y_{1..} = 2166,$$

$$R_{I1} = 486 + 288 + 512 + 678 + 436 + 2884 = 5284, \quad C_{J1} = 2166,$$

$$Y_{I1J1} = 98 + 59 + 111 + 148 + 96 + 645 = 1157$$

より、度数は

$$\frac{678 \cdot 2166 \cdot 1157}{5284 \cdot 2166} = 148.5$$

と最尤推定される。

このモデルの適合度の χ^2 値は

$$\begin{aligned} \chi^2 &= \frac{(148 - 148.5)^2}{148.5} + \frac{(444 - 452.4)^2}{452.4} + \cdots + \frac{(236 - 224.5)^2}{224.5} \\ &= 8.68 \end{aligned}$$

のように減少し、独立モデルの場合の $\chi^2 = 96.39$ と比べ、その90.1%をただ一つのパラメータで説明できることになる。なお、両者の差

$$96.39 - 8.68 = 87.71$$

はプールした2次元分割表(表5.9)の独立モデルに対する適合度の χ^2 統計量

表5.9 均一な群にまとめた2次元分割表：ガンの重症度

$I \backslash J$	軽症 {1}	中症・重症 {2, 3}	
「無職」他 {10, 5, 4, 8}	1,009	5,615	6,624
「技能工」他 {7, 9, 2, 1, 6, 3}	1,157	4,127	5,284
	2,166	9,742	11,908

原表5.7を集約・簡単化したもの。各分類内部では一様化されている。

$$\chi^2 = \frac{11908 \cdot (1009 \cdot 4127 - 5615 \cdot 1157)^2}{6624 \cdot 5284 \cdot 2166 \cdot 9742}$$

$$= 87.71$$

から、原表の χ^2 値 96.39 をほぼ再現している。

の値に等しい。つまり、原表での大きな χ^2 値がこのように簡略化した表で再現され、データの解釈を容易にするのである。

以上より企業群10, 5, 4, 8が企業群7, 9, 2, 1, 6, 3に比べ中・重症度の割合が大きいというはっきりした傾向の見られることがわかった。

ただし、このデータ取得後十数年経過しており、この間の医療環境の著しい変化から現状はこれとかなり異なるだろうことを注意しておこう。

5.3.3 順序カテゴリのモデル化

表5.7では列のカテゴリに自然な順序がある。このような場合、本来は順序を考慮した解析をすべきである。すなわち、ある2群を比較するにしても単に列の出現パターンの異同をいうのではなく、一方が一方に比し相対的に右あるいは左寄りの応答が多いという傾向を主に検出すべきである。それが工程の、あるいは一般に処理の良し悪しを比べることになるからである。

その目的で χ^2 適合度統計量を修正した累積 χ^2 統計量 χ^{*2} が提案されている*。

順序カテゴリに対するもう一つのアプローチは、背景に正規分布やロジスティック分布のような連続分布を想定し、それがいくつかの区分点で区切られた頻度データとして観測されたとしてモデル化するものである。この場合、処理の違いは想定した連続分布の平均や分散の違いとしてとらえられる。

例5.8 〈官能検査〉 表5.10は5種類の食品の官能検査の結果である**。

このデータは、従来、ロジスティック分布などによって解析されているが、前節の方法で行列を分類した結果は

行(食品)は3群; $I_1 = \{1, 2\}$, $I_2 = \{3, 4\}$, $I_3 = \{5\}$

列(評価)は2群; $J_1 = \{1, 2\}$, $J_2 = \{3, 4, 5\}$

となり、食品‘5’が右寄り傾向、‘3’, ‘4’が左寄り傾向、‘1’, ‘2’がその中庸といつた傾向がよくとらえられている。

*) 卷末広津(1983).

**) 卷末広津(1992).

表5.10 官能検査データ

食品 \ 評価	1	2	3	4	5	位置母数	尺度母数
1	9	5	9	13	4	0.07	1.25
2	7	3	10	20	4	0.56	1.05
3	14	13	6	7	0	-1.11	0.90
4	11	15	3	5	8	-0.50	1.66
5	0	2	10	30	2	0.98	0.51

注) 1; 'Terrible', 5; 'Excellent' と評価される (出典: ブラッドレー他)

ここでいう官能検査 sensory test はこの表から想像されるように、試料の品質を感覚によって、あるいは逆に感覚を試料によって測定する手段をさす。

なお、表の右端に示したのはロジスティック分布(4.4)をあてはめたときの位置母数 $-\beta_0/\beta_1$ 、尺度母数 β_1 の推定値である。位置母数が平均的な右寄りあるいは左寄り傾向を表し、尺度母数がデータの拡がりぐあいを表すと考えるとこのあてはめも比較的よくデータを説明している。

表5.3(期末試験データ)ではさらに行、列ともに順序がある。この場合も順位相関をはじめ様々な解析法が提案されている。なお、表5.3(試験データ)および表5.7(ガンの重症度のデータ)はデータの傾向がはっきりしているため、先の結論は、かかる順序を考慮した解析を行っても変わらない。より微妙なデータでは、順序を考慮した解析により、より顕著に傾向がとらえられることがある。

5.4 分割表の対称性の検定

5.4.1 対称性

まず、基本である対称性を例で示そう。

例5.9 <投票パターン> 表5.11は無作為に選ばれた $n=175$ 人のパネラーの2度の総選挙における投票パターンを示したものである。このようなデータでは対角線にデータが集中するのは自明であり、前節までの独立性の検定は意味がない。それに替って興味があるのは支持政党の推移であり、帰無仮説

$$(5.25) \quad H_0: p_{ij} = p_{ji}, \quad 1 \leq i < j \leq a$$

❀ここ一番の大切さ❀

スポーツの対戦成績から選手の強さを決める考え方には幾通りもある。相撲のように2人の力士が何度も対戦するとき、両者の勝敗の比がほぼその「強さ」の比になると想定しよう。例えば、強さがそれぞれ10, 20である力士A, Bが対戦すると、Aの勝率はほぼ1/3となるものと想定する(後の5.5節に述べるブラッドリー・テリーのモデル)。

簡単のため力士数を3人とする。力士A, Bの他に強さが40の力士Cがいるものとし、A, Bが6回、A, Cが5回対戦したとしよう。Aの勝敗は

$$6 \cdot (1/3) + 5 \cdot (1/5) = 3$$

となる。このように各力士の強さとお互いの対戦数から全力士の勝敗が計算できる。

そこで今度は、全力士の勝敗の計算値が実際の勝敗にぴったり一致するような「強さ」を求め、それを各力士の強さとする。これは、必ずしも勝率の順序と一致しない。次の表をみてみよう。

	1位	2位	3位
昭和50	北の湖 227	三重海 140	貴ノ花 126
51	輪島 284	北の湖 211	旭国 130
52	北の湖 458	輪島 281	若乃花 128
53	北の湖 513	若乃花 344	輪島 161
54	北の湖 325	三重海 249	若乃花 200
55	北の湖 289	若乃花 201	輪島 147
56	千代富 280	北の湖 242	琴風 119
57	千代富 280	北の湖 188	隆の里 164
58	千代富 432	隆の里 275	琴風 164
59	若島津 230	千代富 207	隆の里 151
60	千代富 395	双羽黒 142	北天佑 136
61	千代富 398	双羽黒 184	北勝海 130
62	千代富 272	北勝海 264	大乃国 179
63	千代富 702	旭富士 233	北勝海 158

昭和の各年次の強豪力士と強さ(年間6場所制の昭和33年以降のうち昭和50年以後のもの)

(藤野和建「横綱の強さ」UP 233号、1992.3)

各年度の最強力士のうち59年の若島津だけが横綱になっていない。この年、大関若島津は3月場所と7月場所で優勝した。その時点で、過去3場所の成績を通算した強さは344で、他の横綱昇進のケースと比べて全く遜色のないものであった。しかし、9月場所14日目、平幕の多賀竜に敗れ、優勝と横綱の可能性を失う。相撲におけるここ一番の重要さが感じられる。

表5.11[†] 英国における投票の推移を表すデータ

	1970年	1. 保守党	2. 労働党	3. 自由党	計
1966年					
1. 保守党		68	1	1	70
2. 労働党		12	60	5	77
3. 自由党		12	3	13	28
計		92	64	19	175

保守党、労働党、自由党から、次の選挙でどのように票が流れたかを示すデータである（「パネラー」とは一定の判断を行うために構成された、判定者、被験者、検査者などの集団をいう）。対角線が大きいのは当然として、 i 党 → j 党 ($i \neq j$)への流れが重要である。この流れは 2 つの周辺分布 (92, 64, 19) と (70, 77, 28) の変化だけではとらえられないことに注意する（5.4.1 参照）。

の検定に興味が持たれる。これは正方な分割表における対称性の適合度検定 goodness-of-fit test of symmetry とよばれている。

前章の(4.5)において $p_{ij} = p_{ji}$ と置き、最尤推定量を求める

$$(5.26) \quad \bar{p}_{ij} = y_{ij}/n, \quad \bar{p}_{ji} = \bar{p}_{ji} = (y_{ij} + y_{ji})/(2n)$$

が得られ、これから適合度の χ^2 統計量

$$(5.27) \quad \begin{aligned} \chi^2 &= \sum_i \sum_j \left(y_{ij} - \frac{y_{ij} + y_{ji}}{2} \right)^2 / \left(\frac{y_{ij} + y_{ji}}{2} \right) \\ &= \sum_{i < j} (y_{ij} - y_{ji})^2 / (y_{ij} + y_{ji}) \end{aligned}$$

が得られる。

たとえば、例 5.9 の英国の投票パターンのデータ（表 5.11）に対しては、 \bar{p}_{ij} は下表のようになる。

	1970 1966	保守党	労働党	自由党
保守党		0.389	0.037	0.037
労働党		0.037	0.343	0.023
自由党		0.037	0.023	0.074

これに類するものとして次頁の上に示すような左右の裸眼視力を 4 等級に分類したデータなどもよく知られている。

自由度は一般の a カテゴリーの場合に

$$\nu = a^2 - 1 - \left\{ \frac{a(a-1)}{2} - 1 \right\} = {}_a C_2$$

となる。これは帰無仮説(5.25)の個数(線形独立な数、ランク)に等しい。

5.4 分割表の対称性の検定

等級	[左 眼]				
	最高	第2級	第3級	最低	計
右	最高	1520	266	124	66
	第2級	234	1512	432	78
	第3級	117	362	1772	205
	最低	36	82	179	492
	計	1907	2222	2507	841
					7477

(出典: Biometrika, 40)

表[†] 30-39 歳の女性(7477 人)の裸眼視力、1943-6 年、英國軍需工場におけるデータである。左、右とも 4 等級、同一カテゴリに分類されているが、左右の対称性あるいは、2 つの周辺分布が等しいか否かの仮説が問題になる。

例 5.9 の表 5.11 の英國の投票パターンのデータについて、対称性の適合度検定を行ってみよう。 χ^2 および自由度の計算は次のようになる。

$$\chi^2 = \frac{(1-12)^2}{1+12} + \frac{(1-12)^2}{1+12} + \frac{(5-3)^2}{5+3} = 19.2, \quad \nu = {}_3 C_2 = 3.$$

実現値 19.2 は自由度 3 の χ^2 分布 $\chi^2(3)$ の上側 3% 点に当たり、対称性の仮説は棄却される。

5.4.2 周辺対称性と準対称性

対称性の仮説(5.25)は次の 2 つの仮説に分解される。

$$(5.28) \quad \text{周辺対称性 } H_{01}: \quad p_{ij} = p_{ji}, \quad i=1, \dots, a$$

$$(5.29) \quad \text{準対称性 } H_{02}: \quad p_{ij}p_{ja}p_{ai} = p_{ji}p_{ia}p_{aj}, \quad 1 \leq i < j < a$$

仮説 H_{01} の自由度は、(5.28) の式の数から $p_{..} \equiv 1$ により 1 だけ減じて

$$(5.30) \quad \nu_{01} = a - 1,$$

H_{02} については

$$(5.31) \quad \nu_{02} = {}_a C_2 - (a - 1) = {}_{a-1} C_2$$

となる。

(5.25) が成り立てば(5.28), (5.29) が成り立つことは容易に確かめられる。逆も成り立つが証明はやや面倒である。(5.28) の意味は自明であるが(5.29) は少し説明を要する。そのため表 5.12 を見てほしい。

いま保守党から労働党に、労働党から自由党に、そして自由党から保守党に同程度の推移があったとする、各党の総得票数には変化がなく((5.28) は保

表5.12 周辺を動かさない投票パターンの推移(例5.9, 表5.11)

1966 投票		保守 党	労 働 党	自 由 党
1970 投票				
保 守 党	-Δ	Δ	0	
労 働 党	0	-Δ	Δ	
自 由 党	Δ	0	-Δ	

保守党→労働党→自由党→保守党とある得票量(Δと表している)ずつ動いているにもかかわらず、周辺分布には何の動きもないというパターン。つまり、保守党から次の選挙でΔだけ労働党へ流出するが、自由党からの流入がΔだけある等々。

ったまま), 対称性はくずれる。このような意味の非対称性がないとするのが(5.29)の仮説である。なお、(5.29)はすべての3つ組について

$$p_{ij}p_{jk}p_{ki} = p_{ji}p_{ik}p_{kj}$$

が成り立つことと同値であることに注意する。

表5.11(英国の投票のデータ)で(5.28)と(5.29)を別々に検討すると、(5.28)に対してのみ適合度の χ^2 値が1%有意になる。つまり、労働党および自由党から保守党への平均的な推移が認められ、複雑な‘三すくみ’の推移は認められない。このうち、三すくみがないこと(つまり準対称性)の検定は次節のプラッドリー・テリーのモデルの適合度検定に帰着する。

5.5 プラッドリー・テリーのモデル

これは「強さ」を測るモデルである。親しみ深い一例をあげよう。

例5.10 〈勝敗のデータのモデル〉 表5.13は1981年のパ・リーグ公式戦の勝敗表である。この表で y_{ij} はチーム*i*がチーム*j*に勝った回数を表す。いま、チーム*i*がチーム*j*に勝つ確率を p_{ij} と表し、引き分けはないものとすると y_{ij} は、二項分布

$$P(Y_{ij}=y_{ij}) = {}_{y_{ij}+y_{ji}}C_{y_{ij}} (p_{ij})^{y_{ij}} (p_{ji})^{y_{ji}} \quad (p_{ij} + p_{ji} = 1)$$

に従うと考えられる。したがって、表5.13全体では、すべての組合せについての積

$$(5.32) \quad f_p(\mathbf{y}) = \prod_{1 \leq i < j \leq 6} \prod_{(y_{ij}, y_{ji})} C_{y_{ij}} (p_{ij})^{y_{ij}} (p_{ji})^{y_{ji}}$$

5.5 プラッドリー・テリーのモデル

表5.13 1981年パ・リーグ勝敗表(引き分けは省略)

チ ー ム	1(Fi.)	2(Br.)	3(Li.)	4(Or.)	5(Ha.)	6(Bu.)	総勝数
1. ファイターズ (Fi.)	—	14	7	13	16	18	68
2. ブレーブス (Br.)	12	—	15	8	16	17	68
3. ライオンズ (Li.)	16	9	—	12	14	12	63
4. オリオンズ (Or.)	12	17	12	—	13	7	61
5. ホークス (Ha.)	10	10	11	12	—	11	54
6. バッファロウズ (Bu.)	4	8	12	16	13	—	53
総 負 数	54	58	57	61	72	65	367

(出典: 広津(1983))

スポーツは統計データの宝庫であり、いろいろな分析モデルによってさまざまな興味深い結論を与える。このデータも「プラッドリー・テリーのモデル」をあてはめると、各チームの強さ(合計=100とする)として

$$Fi.: 20.0, \quad Br.: 18.8, \quad Li.: 17.9, \quad Or.: 16.5, \quad Ha.: 13.0, \quad Bu.: 13.9$$

という結果を与える。また $>$ で‘かもる’を表すとすると

$$Bu. > \{Li., Or.\} > \{Fi., Br.\} > Bu.$$

という三すくみ(なめくじ>蛇>蛙>なめくじの関係)が見られる。

が確率分布になる。 \mathbf{p}, \mathbf{y} はすべての p_{ij} と y_{ij} のベクトルである。

この場合に勝敗の確率を記述するモデルとしてよく用いられるのがプラッドリー・テリーのモデル Bradley-Terry's model

$$(5.33) \quad p_{ij} = \frac{\theta_i}{\theta_i + \theta_j}, \quad p_{ji} = \frac{\theta_j}{\theta_i + \theta_j}$$

である。これは各チームには強さを表すある正数 $\theta_i (i=1, \dots, 6)$ が対応しており、ある2チーム*i, j*が対戦したときの勝敗の確率が(5.33)の形式で表されるというのである。これはチーム力が1次元的に表現できることを仮定しており、もし成立すれば大変便利な表現である。ところで(5.33)が満たされれば

$$(5.34) \quad p_{ij}p_{jk}p_{ki} = \frac{\theta_i \theta_j \theta_k}{(\theta_i + \theta_j)(\theta_j + \theta_k)(\theta_k + \theta_i)} \\ = p_{ji}p_{ik}p_{kj}$$

が成り立つ。 $p_{ij}p_{jk}p_{ki} > p_{ji}p_{ik}p_{kj}$ のとき*i*が*j*に強く、*j*が*k*に強く、*k*が*i*に強いという三すくみが生じるから、プラッドリー・テリーのモデルは三すくみが無い(したがって1次元に並べられる)モデルといつぐことができる。

このモデルの適合度検定は確率分布(5.32)と(4.5)との違いにもかかわらず、(5.29)と(5.33)の同等性により(5.29)の検定と全く同じ統計量が用いられる。

ここで念のため、具体的にモデルの最尤推定、適合度統計量を紹介しよう。いま、(5.32)を尤度 $L(\mathbf{p})$ としこれに(5.33)を代入し整理すると、対数尤度は

$$(5.35) \quad \log L(\mathbf{p}) = \sum_{1 \leq i \leq j \leq a} \{\log r_{ij} C_{y_{ij}} - r_{ij} \log (\theta_i + \theta_j)\} \sum_i y_{ii} \log \theta_i$$

となる。ただしチーム数を a とし、 i と j の対戦数を $r_{ij} = y_{ij} + y_{ji}$, i の総勝ち数を $y_{ii} = \sum_{j \neq i} y_{ij}$ とおいた。この式を θ_i で偏微分して 0 とおくと、

$$(5.36) \quad \hat{\theta}_i \sum_{j \neq i} \frac{r_{ij}}{\hat{\theta}_i + \hat{\theta}_j} = y_{ii}$$

が得られる。ここで θ_i は相対的な大きさしか意味がなく、実際(5.36)からも比しか定まらない。そこで一般性を失うことなく K を適当な定数として

$$(5.37) \quad \sum_{i=1}^a \hat{\theta}_i = K$$

という式を付加して(5.36)を解くことができる(たとえば $K=100$)。実際には、次のような反復法で解くことになる。いまある段階での近似値を $\hat{\theta}_i^0$ ($i=1, \dots, a$) とする。このとき(5.36)による近似値

$$\hat{\theta}_i = y_{ii} / \sum_{j \neq i} \frac{r_{ij}}{\hat{\theta}_i^0 + \hat{\theta}_j^0}$$

を(5.37)を満たすようにスケーリングして、次段階の近似値

$$\hat{\theta}_i^1 = \frac{K \hat{\theta}_i}{\sum \hat{\theta}_i}$$

とし、この手順を収束するまで繰り返す。

適合度 χ^2 の自由度は(5.31)より $a-1 C_2$ であるが、これは本来 $a C_2$ 通りの 2 項確率 p_{ij} ($1 \leq j < i \leq n$) のうち、ブラッドリー・テリーのモデルには実質的に $a-1$ 個のパラメータ(相対的な比率だけが問題なので)が残されていることからも理解される。

例 5.10 の表 5.13 のパ・リーグの勝敗データに対しブラッドリー・テリーのモデルの適合度検定を行おう。 $K=100$ として(5.36)と(5.37)に基づく反復法を適用すると、「強さ」 $\hat{\theta}$ は次のようになる(なお、一次元でないことがわかる)。

表 5.14 ブラッドリー・テリーのモデルのあてはめ: 推定された勝ち数

チーム	1(Fi.)	2(Br.)	3(Li.)	4(Or.)	5(Ha.)	6(Bu.)
1. ファイターズ (Fi.)	—	13.398	12.149	13.697	15.763	12.992
2. ブレーブス (Br.)	12.602	—	12.310	13.317	15.380	14.391
3. ライオンズ (Li.)	10.851	11.690	—	12.474	14.475	13.511
4. オリオンズ (Or.)	11.303	11.683	11.526	—	13.990	12.498
5. ホークス (Ha.)	10.237	10.620	10.525	11.010	—	11.608
6. バッファロウズ (Bu.)	9.008	10.609	10.489	10.502	12.392	—

(出典: 広津(1983))

各チームの推定された強さから計算された勝ち数、負け数である。ただし、強さは 1 次元的である(三すくみはない)ことを仮定している。これともとの表のデータ(表 5.13)を比較すると、相当に様相が異なることから、強さはこの場合単純な 1 次元のものでないことがわかり、スポーツの勝敗の奥行の深さが示唆される。

$$\frac{\text{ファイターズ}}{\hat{\theta}_1 = 19.993} \cdot \frac{\text{ブレーブス}}{\hat{\theta}_2 = 18.805} \cdot \frac{\text{ライオンズ}}{\hat{\theta}_3 = 17.857} \cdot \frac{\text{オリオンズ}}{\hat{\theta}_4 = 16.499} \cdot \frac{\text{ホークス}}{\hat{\theta}_5 = 12.984} \cdot \frac{\text{バッファロウズ}}{\hat{\theta}_6 = 13.862}$$

これから、たとえばチーム 1(ファイターズ)が 2(ブレーブス)に勝つ確率は

$$\tilde{p}_{12} = \frac{\hat{\theta}_1}{\hat{\theta}_1 + \hat{\theta}_2} = \frac{19.903}{19.993 + 18.805} = 0.5153$$

のように推定される。また、チーム 1 の 2 に対する勝ち数の最尤推定量は

$$\hat{y}_{12} = 26 \cdot 0.5153 = 13.398$$

負け数(チーム 2 の 1 に対する勝ち数)の最尤推定量は

$$\hat{y}_{21} = 26 \cdot (1 - 0.5153) = 12.602$$

のように得られる。その他の組合せに対しても同様に求めた結果は表 5.14 のようになる。これから適合度 χ^2 を計算すると

$$\begin{aligned} \chi^2 &= \frac{(14 - 13.398)^2}{13.398} + \frac{(7 - 12.149)^2}{12.149} + \dots + \frac{(13 - 12.392)^2}{12.392} \\ &= 22.39 \end{aligned}$$

となる。一方、自由度は $\nu = 5 C_2 = 10$ である。 $\chi^2_{0.01}(10) = 23.21$ であることから、実現値 22.39 はほぼ 1% 有意である。つまりブラッドリー・テリーのモデルは適合しない。

例 5.10 の勝敗表(表 5.13)をよく見るとバッファロウズが平均的な勝率に比して上位チーム 1, 2, に「かも」にされたことが見てとれる。そこで j が i を「苦手」とすることを $i > j$ と表すと

$$6 > (3, 4) > (1, 2) > 6$$

という三すくみが見られる。

なお表 5.13 のようなデータは例 5.10 のようなスポーツのリーグ戦のみならず官能検査の分野における一対比較法でよく得られる。たとえば、食品の善し悪しを比べるのに表 5.10 のような等級付けを行うのではなく、何人かのパネラーに各 2 つずつを比べ、どちらを好むかを答えてもらうとまさに表 5.13 と同じ形式のデータになる。そこでもしプラッドリー・テリーのモデルが適合すると、食品を好ましさの順に 1 列に並べ、 $\hat{\theta}_i$ で相対的な‘好まれ度’を推定することができる。

前節の準対称性の仮説 H_{02} の適合度検定を本節の方法で行おうとすると、対角のデータ y_{ii} の処置に惑うかも知れない。しかし、実は H_{02} の検定の際、 y_{ii} は実質的な役割を果たさないことから、本来非対角要素しかない場合のプラッドリー・テリーモデルの適合度検定統計量をそのまま使って差し支えないものである。そもそも対称性仮説検定の際に、仮説を課す、課さないにかかわらず対角要素の推定値は等しく(5.26)のように $\hat{p}_{ii} = y_{ii}/n$ であったことを思い出そう。

5.6 3 次元分割表と対数線形モデル*

同じように 2 次元分割表にまとめられるデータにも実は標本の抽出法(サンプリング法)の違いにより基礎となる確率分布はいろいろである。それにもかかわらず、たとえば行・列の独立性と多項分布の一様性、あるいは正方分割表における準対称性とプラッドリー・テリーのモデルの適合度検定は同じ統計量に帰着した。3 次元分割表(3 重分類表)にはさらにいろいろな場合があるが、ここでも同じようなことが生じる。

それらを統一的に議論するのに対数線形モデル log-linear model が便利である。3 次元分割表にまとめられるデータの代表的なものに、次の 3 通りの種類がある。ここで特に応答変数をスクリプト体(花文字) $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$ で表し、要因と区別することにする。

(i) 2 要因 A, B の各水準組合せでカテゴリカルな応答 \mathcal{C} (多項分布)が観

*) 本節は簡略化した議論である。

$$6 > (3, 4) > (1, 2) > 6$$

表 5.15 混紡洋服地の評価

要因組合せ	評価(風合)			計
	上	中	下	
$A_1 B_1$	7	6	1	14
$A_1 B_2$	3	8	3	14
$A_2 B_1$	7	7	0	14
$A_2 B_2$	3	9	2	14
$A_3 B_1$	4	7	3	14
$A_3 B_2$	1	7	6	14
$A_4 B_1$	5	7	2	14
$A_4 B_2$	0	6	8	14

(出典：田口『統計解析』丸善)

A = 水準組合せ(混紡系の 4 種類),

B = 水準組合せ(仕上げ法の 2 種類),

C = カテゴリカル応答(順序尺度 上, 中, 下)

測されている場合。

これは、2 元配置要因実験でデータが計量値ではなくカテゴリカル応答で得られたと考えればよい。この場合、カテゴリは名義尺度(属性分類)のこともあるが、むしろ表 5.15 のような順序尺度であることが多い*). 製品を良・不良で分ける 2 項分類データや、順位データなどは、その特別な場合と考えることができる。

(ii) 要因 A の各水準で 2 変数カテゴリカル応答(\mathcal{B}, \mathcal{C})(2 次元分割表)が観測されている場合。

たとえば、表 5.3 のようなデータが別の年度にも得られているとすると、年度によって代数(\mathcal{B})と解析(\mathcal{C})の成積の相関が異なるか否かに興味が持たれる。これは年度を要因 A として $A \times \mathcal{B} \times \mathcal{C}$ の例を与える。

(iii) 3 変数カテゴリカル応答($\mathcal{A}, \mathcal{B}, \mathcal{C}$)(3 次元分割表)が観測されている場合。

たとえば、表 5.3 で代数、解析に加えて幾何の成積も得られていると、それは、3 次元表 $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$ の例を与える。

さて、これら 3 通りの場合を対数線形モデル

$$(5.38) \quad \log p_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk}$$

を用いることにより、まったく同じ形式で扱うことができる。ここで分散分析

*) 本シリーズ I 「統計学入門」の尺度の項(p. 27)参照。

モデルの場合と同じように $\alpha_i, \beta_j, \gamma_k$ を主効果、 $(\alpha\beta)_{ij}, (\alpha\gamma)_{ik}, (\beta\gamma)_{jk}$ を2因子交互作用、 $(\alpha\beta\gamma)_{ijk}$ を3因子交互作用とよぶ。 $(\alpha\beta)_{ij}$ は2因子の組合せ効果、 $(\alpha\beta\gamma)_{ijk}$ はそれが他の因子の水準に応じていかに異なるかを表すが、その実際的な解釈は上記(i), (ii), (iii)に対応してそれぞれ異なる。たとえば三因子交互作用 $(\alpha\beta\gamma)_{ijk}$ は、(i)では要因 A, B の交互作用が応答 C に影響を与えるか否か(つまり、2元配置分散分析モデルと同じ意味での2因子交互作用)を表し、(ii)では2変数 B, C の関連の度合が要因 A の各水準で異なるか(その意味で A の効果があるか)否かを表し、(iii)では3変数 A, B, C の関連の度合を表す。このような解釈に気をつけさえすれば、これら3通りの場合が同様に扱えるといふのはたいへん便利な性質である。

さらに、4次元、5次元分割表のデータにもよく出あう。それらでは要因と応答変数の取り合せにさらに様々な場合が生じるが、それらも対数線形モデルを考えることにより同一に扱える。

なお、直観的には p_{ijk} に直接(5.38)の右辺のような構造を仮定する方がわかりやすい。しかし、それでは次に述べる「シンプソンのパラドックス」のような不合理が避けられず、長い論争の末、対数線形模型が多く使われるようになったものである。

解析手法としては χ^2 適合度検定よりむしろ尤度比検定を用い、高次の交互作用から順に検定して適合するモデルを決定する。

ただし、次元数が大きく検定を繰り返すような場合には、重回帰分析における変数選択手法、たとえば AIC 規準などを用いることもある。

最後に対数線形モデルがよく使われるきっかけとなった、シンプソンのパラドックス Simpson's paradox を紹介しておこう。

例 5.11 〈シンプソンのパラドックス：対数線形モデル〉 表 5.16において老年、若年の層ごとに見ると薬効に差はなく、したがって3因子交互作用もない。実際、有効対無効の比(オッズという)は、薬剤 1, 2 に対しそれぞれ老年では 40/10, 120/30、また老年では 30/120, 10/40 となって各層ごとに等しい。このような場合、通常の加法モデルだと層を併合することが正当化されるが、それだと併合した表に見るように、薬剤 2 が薬剤 1 に比べてはるかに有効という結論が導かれてしまう。

表 5.16 シンプソンのパラドックスの数値例(層別した薬効データ)

原 表	層 別(i)	薬剤(j)	効果の有無(k)		
			1. 有効	2. 無効	
1. 若年		1	40	10	
		2	120	30	
2. 老年		1	30	120	
		2	10	40	
併合(プール)した表		1	70	130	
		2	130	70	

年齢という因子で層別した3次元の分割表の例である。

正しい解釈は、薬剤によらず老年より若年の方が治ゆ率が高く、その若年で薬剤 2 が多く使われた結果、見掛け上薬剤 2 が 1 より有効に見えたということだろう。これをもっと極端にすると、それぞれの層では薬剤 2 が優れ、かつ3因子交互作用が無いのに、層を併合した表では薬剤 1 が見掛け上有意に優れてしまうということも起こり得る。これが「パラドックス」といわれるわけである。

対数線形モデルの場合は、たとえ3因子交互作用が無くてもただちに層を併合することにはならずパラドックスは生じない。これが対数線形モデルがよく使われるようになった理由である。

◆◆◆◆ 第5章 練習問題 ◆◆◆◆

5.1 $\langle \chi^2$ 適合度検定統計量の分布：近似の問題 \rangle 多項分布の χ^2 適合度検定統計量の分布は、セルの期待度数が小さいとき χ^2 分布で必ずしもよく近似できない。そのような例として、 $n=9, p_1=p_2=p_3=1/3$ の3項分布

$$p(x_1, x_2, x_3) = \frac{9!}{x_1! x_2! x_3!} \cdot \left(\frac{1}{3}\right)^9 \quad (*)$$

で、 $(x_1, x_2, x_3)=(6, 2, 1)$ が観測された場合を考えよう。この場合、

$$\chi^2 = (6-3)^2/3 + (2-3)^2/3 + (1-3)^2/3 = 14/3 = 4.67$$

であり、 $\chi^2_{0.10}(2)=4.61$ だから、 χ^2 近似によれば、4.67 以上の値が得られる確率は約 10% となる。

近似の良不良を見るためにこの確率を正確に求めてみよう。いま、 $n=9$ を 3 つに分割する型は全部で 12 通りあるが、このうち χ^2 が 14/3 以上になるのは下記の 7 つの型である。たとえば一番上の 9, 0, 0 の確率は (*) から $1/3^9$ であり、そのような型

を与える観測値は $(x_1, x_2, x_3) = (9, 0, 0), (0, 9, 0), (0, 0, 9)$ の 3 個だから、型 $9, 0, 0$ の確率は $3/3^9$ となる。他の型についての空欄を埋めて完成し、 $\chi^2 \geq 14/3$ となる確率が $3261/19683 = 0.166$ となることを確かめよ。

分割の型	確 率	個 数	型の確率	χ^2
9, 0, 0	$1/3^9$	3	$3/3^9$	18
8, 1, 0				
7, 2, 0				
7, 1, 1				
6, 3, 0				
6, 2, 1				
5, 4, 0				

5.2[†] 〈擬似乱数の適合度検定〉 次表はある方法(ボックス・ミュラー法(第11章参照))によって生成した 40000 個の擬似標準正規乱数の度数分布である。この乱数が標準正規分布に従うという仮説の χ^2 適合度検定を行え(有意水準 5%)。

区 間	度 数
~ -3.0	60
-3.0 ~ -2.5	208
-2.5 ~ -2.0	655
-2.0 ~ -1.5	1826
-1.5 ~ -1.0	3650
-1.0 ~ -0.5	5909
-0.5 ~ 0.0	7512
0.0 ~ 0.5	7737
0.5 ~ 1.0	6063
1.0 ~ 1.5	3641
1.5 ~ 2.0	1820
2.0 ~ 2.5	683
2.5 ~ 3.0	188
3.0 ~	48
計	40000

5.3[†] 〈2×4 分割表の検定〉 ドラム缶の巻締不良を 4 人の作業員別に集計して、次の結果を得た。 χ^2 適合度検定を用い、作業員によって不良率に差がないという仮説を有意水準 5% で検定せよ。

作業員 良, 不良	A	B	C	D	計
良	7142	3021	1841	5880	17884
不 良	90	60	38	29	217
計	7232	3081	1879	5909	18101

(出典:『品質管理便覧』)

5.4 〈適合度検定: 試合数の分布〉 (i) プロ野球日本シリーズでは最初に 4 勝したチームが優勝する。両チームが互角で、試合結果が毎回独立なら、引き分けを除くシリーズの試合数 z の分布は

$$p(z) = {}_{z-1}C_3 2^{-z+1}, \quad z=4, 5, 6, 7$$

で与えられることを示せ。

(ii) 次表は 1950 年から 1991 年まで 42 年間の日本シリーズの試合数を示す。試合数が(i)の分布に従うという仮説の χ^2 適合度検定を有意水準 5% で行え。

試合数	4	5	6	7	計
度 数	5	8	15	14	42

5.5 〈同: 得点データへの分布のあてはめ〉 次表は 1991 年度プロ野球の西武対近鉄 26 試合での西武の 1 イニングごとの得点の分布である。

得 点	0	1	2	3	4	5	計
度 数	175	24	16	8	3	2	228

(i) 得点の平均 \bar{x} と不偏分散 s^2 を求めよ。

(ii) 得点に \bar{x} を平均とするポアソン分布をあてはめよ。その際、得点 4 と 5 は期待度数が少ないので、1 つに併合せよ。さらに、このあてはめがよくないことを χ^2 適合度検定を用いて確認せよ。

(iii) ポアソン分布のあてはめがうまくいかないので、負の二項分布

$$p(x) = \binom{q+x-1}{x} p^x (1-p)^{q-x}, \quad x=0, 1, 2, \dots$$

のあてはめを試みる。ポアソン分布は平均と分散が等しいが、負の二項分布は平均よりも分散の方が大きいからである。負の二項分布の母数 p, q の推定法のひとつはモーメント法で、それは

$$\hat{p} = \bar{x}/s^2, \quad \hat{q} = \bar{x}^2/(s^2 - \bar{x})$$

とするものである。これらの推定値を求め、負の二項分布をあてはめよ。あてはめのよさを χ^2 適合度検定を用いてチェックせよ。

(注) 任意の実数 a と非負整数 r に対して

$$\binom{a}{r} = \frac{a(a-1)\cdots(a-r+1)}{r!}$$

と定義する。ただし、 $\binom{a}{0} = 1$ とする。

5.6 〈ブラッドリー・テリーのモデル〉 下表は 1991 年度プロ野球セ・リーグ勝敗表である。 $K=100$ として、本文で述べた方法により各チームの強さ θ を推定し、モデルのあてはめのよさを χ^2 適合度検定を用いてチェックせよ。

	広	中	ヤ	巨	大	阪	勝 数
広 島	一	15	15	15	13	16	74
中 日	11	—	14	15	14	17	71
ヤ ク ル ト	11	12	—	14	17	13	67
巨 人	11	11	12	—	13	19	66
大 洋	13	12	9	13	—	17	64
阪 神	10	9	13	7	9	—	48
敗 数	56	59	63	64	66	82	390

5.7 < π の数字の一様性> 円周率 π の小数点以下の数字の並びの内訳は下表の通りである。小数点以下の数字が等確率で出現するという帰無仮説を各桁数ごとに検定せよ。

桁 数	0	1	2	3	4
10^4 桁	968	1026	1021	974	1012
10^5 桁	9999	10137	9908	10025	9971
10^6 桁	99959	99758	100026	100229	100230
10^7 桁	999940	999333	1000306	999964	1001093
10^8 桁	9999922	10002475	10001092	9998442	10003863
10^9 桁	99993942	99997334	100002410	99986911	100011958

桁 数	5	6	7	8	9
10^4 桁	1046	1021	970	948	1014
10^5 桁	10026	10029	10025	9978	9902
10^6 桁	100359	99548	99800	99985	100106
10^7 桁	1000466	999337	1000207	999814	10000401
10^8 桁	9993478	9999417	9999610	10002180	9999521
10^9 桁	99998885	100010387	99996061	100001839	100000273

(出典：金田康正氏の御好意による)

第6章 検定と標本の大きさ

マウスを使った実験で特定の薬の効果を確かめるようなとき、一体何匹のマウスを使えばはっきりとした結論が得られるのだろうか。このような問題は、第I巻12章で学んだ仮説検定および検出力の考え方を用いて答えることができる。この章では仮説検定における検出力の考え方と、十分な検出力を得るために標本の大きさの決定の方法について説明する。また、最強力検定の概念を紹介する。

6.1 検定の検出力

検定の考え方についてはすでに説明したが^{*)}、ここでは検定の概念について簡単に復習したうえで、さらに検出力関数の考え方について説明しよう。

6.1.1 検定の例と考え方

検定とは確率分布の母数 θ に対する2つの仮説のいずれが正しいかをデータから決めることである。

例 6.1 <不良率> 1つの例として、ある特定の製品(たとえばカセット・レコーダー)を毎日多数製造している工場を考えよう。ここで関心の的となっているのは、この工場での不良率、すなわち全製品中の不良品の割合であるとする。不良品をゼロにすることはなかなか難しいことであるので、この工場ではたとえば不良率が 1% 以下ならばよしとしているでしょう。しかし、不良品が 1% (やや大きすぎるが、簡単のためこの値とする)をこえた場合には、製造工程になんらかの異常があると考え、その原因を追求する必要がある。

*) 本シリーズ I『統計学入門』第12章。

の判断をどんな方式で行うべきであろうか。

p を不良率とし $p_0=0.01$ とすれば、製造工程が正常であるという仮説は

$$H_0: p \leq p_0$$

と表すことができる。また、製造工程が異常であるとする仮説は

$$H_1: p > p_0$$

と表すことができる。 H_0 を帰無仮説 null hypothesis といい H_1 を対立仮説 alternative hypothesis といいう。データに基づいて H_0 が正しいと判断することを帰無仮説を採択または受容 accept するといい、 H_1 が正しいと判断することを帰無仮説を棄却 reject するといいう。

一般には、帰無仮説に対応する母数 θ の集合を Θ_0 とおき、対立仮説に対応する母数 θ の集合を Θ_1 とおけば、帰無仮説を $H_0: \theta \in \Theta_0$ または対立仮説を $H_1: \theta \in \Theta_1$ と表すことができる。上の例では $\Theta_0=[0, 0.01]$, $\Theta_1=(0.01, 1]$ である。

さて上の例で考えた工場では、1日の終わりに $n=100$ 個の製品を抜取検査し、その中の不良品の数 X が $X \geq 3$ の時製造工程に異常があるのではないかと判断して調査をし、 $X=0, 1, 2$ の場合には正常の状態であるとしてそのまま翌日の操業を続けるものとしよう。この場合の検定方式は

(6.1) $X \geq 3$ ならば帰無仮説を棄却、 $X \leq 2$ ならば帰無仮説を採択と表すことができる。この検定方式を δ_1 と表そう^{*}。 X の分布は二項分布 $Bi(100, p)$ であると考えることができるから、この場合二項分布における $\theta=p$ が関心のある未知母数である。さて、(6.1)の検定では 3 個以上の不良品の場合に帰無仮説を棄却したが、たとえば 4 個以上の不良品で帰無仮説を棄却するいま一つの検定方式 δ_2 を考えることもできる。

δ_1 と δ_2 のどちらが望ましい検定方式であろうか。そのためには、それぞれの検定方式を用いたときに誤った判断をする確率を考えればよい。

6.1.2 検定における過誤と検出力

検定において、帰無仮説が正しいときに帰無仮説を棄却してしまう誤りを第

*¹) δ は‘デルタ’と読む。検定(方式)は、 t, ϕ などいろいろな表し方があるが、ここでは A. ワルドの統計的決定理論での表記を用いた。特に混乱はないであろう。

1種の過誤 error of the first kind とよび、対立仮説が正しいときに帰無仮説を採択してしまう誤りを第2種の過誤 error of the second kind とよぶ。 δ_1 を用いた検定においては、 $p \leq 0.01$ であるのに(たまたま) $X \geq 3$ となった場合に第1種の過誤をおかすことになり、 $p > 0.01$ であるのに $X \leq 2$ が観測された場合に第2種の過誤をおかすことになる。良い検定とは第1種の過誤の確率および第2種の過誤の確率のいずれもが小さい検定であると考えられる。このことを次のように検出力関数であらわそう。

検出力関数 いま検定方式 δ を用いたとして、 θ という値に対して帰無仮説を棄却する確率を

$$(6.2) \quad \beta_\delta(\theta) = P_\theta(\delta \text{を用いて帰無仮説を棄却})$$

と表す。 $\beta_\delta(\theta)$ を検出力 power あるいは、 θ の関数として検出力関数 power function といいう。ただし P_θ は θ が真のときの確率を表す(下例参照)。一般に検定方式 δ は観測値 X_1, \dots, X_n の関数 $t=t(X_1, \dots, X_n)$ とある限界値 t_0 を用いて、 $t > t_0$ ならば帰無仮説を棄却し、 $t \leq t_0$ ならば帰無仮説を受容するという不等式で表すことが多い。この場合の t を検定統計量 test statistic といい、限界値 t_0 を棄却点 critical point あるいは有意点 significance point といいう。また $t > t_0$ となる範囲を棄却域 critical region といいう。このとき、検定方式 δ の検出力関数は

$$(6.2') \quad \beta_\delta(\theta) = P_\theta(t(X_1, \dots, X_n) > t_0)$$

と表される。

例 6.2 〈続〉 例 6.1 の δ_1 では $\theta=p$ で $t(X)=X$ であり、 $t_0=2$ である。その検出力関数は、 p の関数として

$$(6.3) \quad \beta_{\delta_1}(p) = P_p(X \geq 3)$$

$$= 1 - P_p(X=0) - P_p(X=1) - P_p(X=2)$$

$$= 1 - (1-p)^n - np(1-p)^{n-1} - \frac{n(n-1)}{2} p^2 (1-p)^{n-2}$$

となる。同じく、 δ_2 については

$$\beta_{\delta_2}(p) = P_p(X \geq 4)$$

$$= \beta_{\delta_1}(p) - P_p(X=3)$$

となっている。

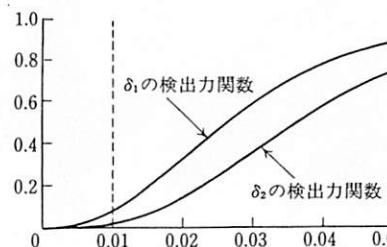


図 6.1 例 5.1 に対する検出力関数のグラフ
 δ_1 の方が検出力が高く見えるが、有意水準($p=0.01$ のときの値)も高くなっているので、優劣は決められない。
 しかし、有意水準さえ一定値にそろえれば、検出力関数は検定の良否、優劣を判断する重要な基準となる。

検出力関数はグラフに描いてみるとわかりやすい。図 6.1 は(6.3)の δ_1 の検出力関数 $\beta_{\delta_1}(p)$ および δ_2 の検出力関数 $\beta_{\delta_2}(p)$ をグラフにしたものである。また表 6.1 では、いくつかの p の値について $\beta_{\delta_1}(p)$ および $\beta_{\delta_2}(p)$ の値を示している。

良い検定方式 検出力関数を用いれば、検定方式の第 1 種の過誤の確率は、棄却の確率だから

$$\beta_{\delta}(\theta), \quad \theta \in \Theta_0$$

また、第 2 種の過誤の確率は、採択の確率だから

$$1 - \beta_{\delta}(\theta), \quad \theta \in \Theta_1$$

と表すことができる。すなわち第 2 種の過誤の確率は対立仮説における検出力を 1 から引いたものである。したがって良い検定方式とは 2 つの過誤の確率が小さいもの、すなわち帰無仮説が正しいときに検出力関数 $\beta_{\delta}(\theta)$ の値が小さく、対立仮説が正しいときに検出力関数の値の大きいような検定方式であることがわかる。

しかし、両方の過誤の確率を同時に小さくすることはできない。上の例の δ_1 では第 1 種の過誤の確率の最大値は $p=0.01$ のときの $\beta_{\delta_1}(0.01) = 0.079$ である。またたとえば $p=0.05$ のときの検出力は $\beta_{\delta_1}(0.05) = 0.882$ 、第 2 種の過誤の確率は 0.118 である。したがって、 $p=0.05$ であり生産工程に異常があるときには、0.882 の確率で帰無

表 6.1 二項分布に対する 2 つの検定方式 δ_1 , δ_2 の検出力およびその比較(標本の大きさ $n=100$ の場合)

p の値	検出力 $\beta_{\delta_1}(p)$	検出力 $\beta_{\delta_2}(p)$
0.01	0.079	0.018
0.02	0.323	0.141
0.03	0.580	0.353
0.04	0.768	0.571
0.05	0.882	0.742

3 以上のときに棄却する検定方式 δ_1 の方が、4 以上のときに棄却する δ_2 よりも検出力が高い。

✿うるさい火災報知器✿

ホテルや旅館で火災が発生すると多くの犠牲者が出るためにニュースになることが多い。このような場合に火災報知器が正しく作動していたかが問題となるが、時には火災報知器のスイッチが切られていたこともあるようである。「火災報知器のスイッチを入れておくと、何でもないのにしょっちゅう火災報知器がなってうるさいので普段は止めておいた」ということらしい。

さて「火事が起きない」という通常の状態を帰無仮説とし、火災報知器がなることを帰無仮説を棄却することであると考えてみよう。この場合「何でもないのに火災報知器がなる」ということは、火災報知器が第 1 種の過誤をおかしていることにあたる。この火災報知器は第 1 種の過誤をおかす確率が高く、そのためにきらわれてスイッチを切られてしまったのであろう。スイッチを切られてしまったので、この火災報知器は第 1 種の過誤をおかす確率は 0 になった。しかしながらこのために、実際に火災が起きたときには火災報知器が役に立たず、第 2 種の過誤をおかすハメになったわけである。

仮説を棄却し生産工程の異常を調べることになる。同じ計算を δ_2 についても行うと、 $\beta_{\delta_2}(0.01) = 0.018$ であり、 $p=0.05$ のときの検出力は $\beta_{\delta_2}(0.05) = 0.742$ 、第 2 種の過誤の確率は 0.258 となっている。また、あきらかにすべての p について $\beta_{\delta_1}(p) \geq \beta_{\delta_2}(p)$ である。したがって、これらのふたつの検定方式を比べると、 δ_1 の第 1 種の過誤の確率は δ_2 のそれより大きいが、第 2 種の過誤の確率については逆になっていることがわかる。したがって、このままではどちらの検定方式が良いかを決めることができない。

一般には、仮説検定では第 1 種の過誤を重視し、通常まず第 1 種の過誤の確率を小さくしようとする。なぜなら、仮説検定のもとの形の有意性検定^{*)}では帰無仮説の検証だけが問題になっており、そのときに冒される第 1 種の過誤から検定方式が決ったからである。すなわち有意水準 level of significance と呼ばれるある小さな数 α (通常は $\alpha=0.05$ あるいは $\alpha=0.01$ とする)を与えておいて、第 1 種の過誤の確率を α 以下に押さえ、

$$(6.4) \quad \beta_{\delta}(\theta) \leq \alpha, \quad \theta \in \Theta_0$$

なる検定方式だけを考える。多くの場合等号=が成立し、帰無仮説において

*) 本シリーズ I『統計学入門』第 12 章。

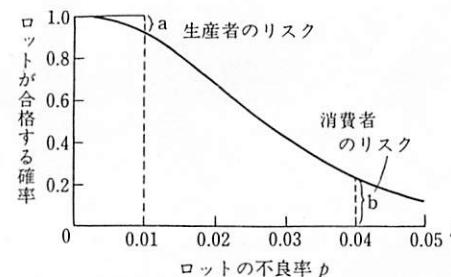


図 6.2 抽取検査の OC 曲線
 $p=0.01$ のときの、図の a の大きさは、誤ってロットが不合格となる確率であり、生産者危険とよばれ、 $p=0.04$ とあるのは、買手が低く要求する不良率の基準値である。 b は、誤って合格とされる確率で、消費者危険といわれる。

$\beta_s(\theta)$ の最大値は有意水準に一致する。 (6.4) を満たす検定方式を有意水準 α の検定方式という。上の例では、 δ_1 は有意水準 5% の検定ではないが、 δ_2 は有意水準 5% の検定となっている。したがって有意水準を 5% とするならば、 δ_2 を使うことになる。

なお、この例では確率変数が離散分布を持つために、第 1 種の過誤の確率の最大値が有意水準 5% に一致してはいないが、連続分布の場合には、棄却点を適切にとることにより第 1 種の過誤の確率の最大値を有意水準に一致させることができるので、そのように棄却点をとるのが普通である。

検出力の問題を、製品の不良率から工程について調べる場合を例にとって説明したが、このような問題は品質管理 quality control における抜取検査に関連している。品質管理の分野では検出力関数のことを作業特性曲線あるいは OC 曲線 operating characteristic curve とよんでいる。ただし、OC 曲線では y 軸を第 2 種の過誤の確率 $1 - \beta_s(\theta)$ とするので検出力曲線とは y 軸の上下が反対になる。図 6.2 は図 6.1 と同じものであるが、検出力曲線を OC 曲線の形に書き直したものである。

製品の出荷にあたっての抜取検査では、出荷されるひとまとまりの製品のあつまりをロット lot とよび、その合否の決定が目的となる。第 1 種の過誤をおかすと、不良品の少ない良いロットであるのに不合格とすることとなり生産者にコストがかかる。この意味で第 1 種の過誤の確率を生産者危険 producer's risk といい、また第 2 種の過誤をおかすと、多数の不良品を含む悪いロットとされて出荷され買手が迷惑するので、第 2 種の過誤の確率を消費者危険 consumer's risk といっている。

6.1.3 検定の分類

いろいろな具体的な検定の検出力を整理するための準備として、ここでは検定を分類するために有用な用語や概念について説明しよう。

考えている母集団の確率分布を表す密度関数あるいは確率関数を $f_\theta(x)$ とする。 θ が関心のある母数である。 X_1, \dots, X_n をこの母集団からの標本の大きさ n の標本とする。すなわち X_1, \dots, X_n は $f_\theta(x)$ を密度関数あるいは確率関数として持つ分布に従う独立な確率変数である。いま θ を一次元の実母数とするとき、検定問題は通常

$$(6.5) \quad H_0 : \theta \leq \theta_0, \quad H_1 : \theta > \theta_0$$

あるいは

$$(6.6) \quad H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0$$

のいずれかの形に定式化される。 (6.5) の検定問題を片側検定 one-sided test といい、 (6.6) の検定問題を両側検定 two-sided test という。たとえば前節の製品の不良率に関する例は片側検定問題である。

以上の説明は考えている母集団が一つだけの場合であったが、検定の応用ではふたつの母集団を比較したい場合も多い。男女間にある基準で差があるかどうか、新薬と従来の薬(標準薬)、ないしはプラセボ、の間で効果の差があるかどうか、などの問題がこの例である。ふたつの母集団の比較を行う検定を 2 標本問題 two-sample problem という。それに対して一つの母集団のみに関心がある検定問題を 1 標本問題という。

ふたつの母集団の分布をそれぞれ $f_{\theta_1}(x), f_{\theta_2}(x)$ とし、母集団 1 からの観測値を X_1, \dots, X_m 、母集団 2 からの観測値を Y_1, \dots, Y_n とする。2 標本問題における片側検定問題と両側検定問題はそれぞれ

$$(6.7) \quad \begin{aligned} H_0 : \theta_1 &\leq \theta_2, & H_1 : \theta_1 &> \theta_2 \\ H_0 : \theta_1 &= \theta_2, & H_1 : \theta_1 &\neq \theta_2 \end{aligned}$$

と表される。たとえば、男女の嗜好(量化されているとして)に差がないかどうかを検定するような場合は両側検定問題である。従来の薬に対して新薬の効果を検定するような場合には、 θ_1 を新薬の効果を表す母数、 θ_2 を従来の薬の効果を表す母数とすれば、片側検定の問題となる。

2 標本問題の応用では、対標本 paired sample とそうでない一般の場合を区

別する必要がある。対標本とは、2つの母集団からペアをなす観測値を取り出して、ペアごとに比較する場合である。たとえば、男女についてある量を比較する場合、 n 組みの夫婦(あるいは、兄弟姉妹)を選んで違いをそれぞれ観測するような場合が対標本である。これに対して、上で説明したのは2つの母集団からそれぞれ独立に標本をとった場合であり、一般的2標本問題の場合である。男性 m 人、女性 n 人をそれぞれ独立に抽出して比較するのが一般的2標本問題にあたる。

動物実験で薬の効果を調べる場合などでも、体重、年齢、遺伝的性質等の条件の似た動物をペアにしてどちらか一方に薬をやるという対実験を行うことが多い。これは、対標本を用いることにより、当の条件以外の外的な諸条件をそろえた上で比較することができるためである。対標本 (x_i, y_i) , $i=1, \dots, n$ においては、観測値の差 $d_i = x_i - y_i$, $i=1, \dots, n$ を新しい観測値として、1標本問題の手法を適用すればいいので、統計的方法はむしろ容易である。次のデータもその一例である。

対標本の例：1988年8月の東京と大阪における最高気温(℃)

日付	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
東京 (X)	32.1	26.2	27.5	31.8	32.1	31.2	30.1	32.4	32.3	29.9	29.6	26.6	31.2	30.9	29.3
大阪 (Y)	35.4	34.6	31.1	32.4	33.3	34.7	35.3	34.3	32.1	28.3	33.3	30.5	32.6	33.3	32.2

(出典：本シリーズI『統計学入門』228頁)

さらに3個以上の母集団を比較する検定問題も考えられる。考えている母集団の個数が k 個であればこの場合の検定問題を k 標本問題とよぶ。第3章の分散分析の問題などが k 標本問題の典型的な例である。

k 標本問題の場合の検出力関数の計算は複雑となるので、ここではこれ以上の説明はやめよう。

6.2 主な検定と検出力

検出力の考え方および検定の分類法を説明したので、主な検定を整理し検定問題における検出力関数を求めておこう。簡単な例でその求め方を学んでほしい。

い。ただし、一般には検出力関数を求めるための積分の計算は複雑となることが多い。また、スペースの関係から、考える母集団分布は正規分布および二項分布に限っておく。

6.2.1 正規分布に関する検定

まず、正規分布の $N(\mu, \sigma^2)$ の母数 μ, σ^2 に関する検定について述べる。 X_1, \dots, X_n を互いに独立に正規分布 $N(\mu, \sigma^2)$ に従う観測値とする。

母平均 μ に関する検定

a) $\sigma^2 = \sigma_0^2$ が既知の場合。まず両側検定問題を考え、帰無仮説および対立仮説を $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ とする。検定統計量は標本平均 \bar{X} を期待値と分散で標準化した

$$(6.8) \quad Z = \sqrt{n}(\bar{X} - \mu_0)/\sigma_0$$

である。帰無仮説のもとで Z が標準正規分布に従うことについて注意すれば、 $z_{\alpha/2}$ を標準正規分布の上側 $\alpha/2$ 点として棄却域を

$$(6.9) \quad |Z| > z_{\alpha/2}$$

とすれば有意水準 α の検定が得られる。

この検定 δ の検出力関数を具体的に計算してみると

$$\begin{aligned} (6.10) \quad \beta_\delta(\mu) &= P_\mu(|Z| > z_{\alpha/2}) \\ &= 1 - P_\mu\left(-z_{\alpha/2} < \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} < z_{\alpha/2}\right) \\ &= 1 - P_\mu\left(-z_{\alpha/2} - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0} < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0} < z_{\alpha/2} - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right) \\ &= 1 - \Phi\left(z_{\alpha/2} - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right) + \Phi\left(-z_{\alpha/2} - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right) \end{aligned}$$

と表される。ここで P_μ は μ が真の母数であるときの確率を表し、 $\Phi(z)$ は標準正規分布の累積分布関数を表す。

片側検定 $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ の場合にも、同じ(6.8)の検定統計量 Z を用いて棄却域を

$$(6.11) \quad Z > z_\alpha$$

※日本シリーズの見込み違い※

プロ野球日本一を決める日本シリーズでは先に4勝したチームを勝者としている。しかしながら、試合に運はつきものであるから、必ずしも強い方のチームが勝者となるわけではない。

そこで、強いチームが勝者となる確率を考えてみよう。いま、 A, B の二つのチームが戦うとして A が B に勝つ確率を p とする。 A が B より強いということは $p > 0.5$ と考えることができる。いま簡単のために、各試合の勝ち負けは独立に決まるとすると、日本シリーズ方式で A が勝者となる確率は

$$f(p) = 0.5 + 2.1875(p - 0.5) - 8.75(p - 0.5)^3$$

で近似できることが示される。

たとえば $p=0.6$ としてみよう。このとき日本シリーズで A が勝つ確率は 0.71 程度である。すなわち、ペナントレースで戦ったとした場合に勝率 6 割で勝ち越せる相手でも、日本シリーズでは 29% の確率で負けてしまうことがあり得るわけである。以上の考え方も検出力に似ている。

(出典：竹内啓・藤野和建『スポーツの数理科学』共立出版)

とおけばよい。この場合の検出力関数は

$$(6.12) \quad \beta_s(\mu) = 1 - \Phi\left(z_{\alpha} - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right)$$

で与えられる。

例 6.3 〈検出力の計算例〉 $n=25, \sigma_0=1, \alpha=0.05, \mu=0.4$ の場合の(6.10)および(6.12)の検出力の値を計算してみよう。両側検定および片側検定の棄却点は $z_{0.025}=1.96, z_{0.05}=1.645$ であり、 $\sqrt{n}(\mu - \mu_0)/\sigma_0=2.0$ である。正規分布表より

$$\Phi(-0.04)=0.498, \quad \Phi(-3.96)=0.0002, \quad \Phi(-0.355)=0.361$$

となるから、両側検定、片側検定の検出力の値はそれぞれ 0.516, 0.639 となることがわかる。 $\mu=0.4$ に対して $n=25$ では、検出力はあまり高くないことがわかる。検出力関数のグラフを図 6.3 に示しておいた。

母分散が既知の場合の Z に基づく以上の 2 つの検定は、母分散が未知の場合の t 検定と対比して z 検定とよばれることがある。

b) σ^2 が未知の場合。以上では母分散 σ^2 が既知という仮定のもとで母平均

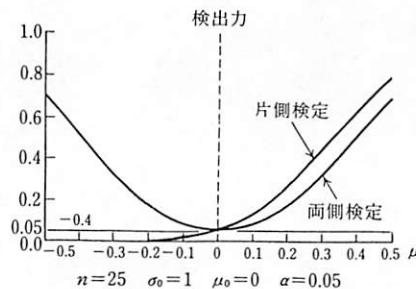


図 6.3 正規分布の(母)平均 μ の検定の検出力
正規分布が仮定された場合の検出力関数の典型である。帰無仮説 $\mu=0$ のところで、有意水準(いまの場合 $\alpha=0.05$)の値に一致する。片側対立仮説が用いられているときは、片側検定が両側検定よりもふさわしいことがわかる。

なお、両側検定の検出力曲線は帰無仮説のところで最小値をとっている(6.4 節の不偏検定参照)。

μ に関する検定を考えたが、 σ^2 が既知という想定は現実的ではない。通常は σ^2 は未知であり、その場合に用いられるのが t 検定 t -test である。いま標本分散を

$$s^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)$$

とする。標本分散は $n-1$ で割ったものとして定義している^{*)}ことに注意する。ここで用いる検定統計量は

$$(6.13) \quad t = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s}$$

である。(6.13)の t は t 統計量 t statistic とよばれる。とくに、回帰係数においては、 t 値、 t 比などともいう。 t 統計量は(6.8)の Z において未知の σ_0 をその推定値 s で置き換えたものになっている。両側検定および片側検定問題の有意水準 α の棄却域はそれぞれ

$$(6.14) \quad |t| > t_{\alpha/2}(n-1), \quad t > t_{\alpha}(n-1)$$

で与えられる。ただし、 $t_{\alpha}(v)$ は自由度 v の t 分布の上側の点、つまり上側確率 = α となる点である。

ところで(6.14)の二つの検定の検出力関数を求めるには、非心 t 分布とよばれる密度関数の積分を必要とし面倒である。しかしながら n が大きいときには、これらの検定の検出力関数は σ^2 が既知の場合の(6.10)および(6.12)の検出力関数で近似されることがわかるので、これらの検出力関数を用いればよい。ただしこの場合、真の母分散 σ_0^2 が未知であることに注意しなければならない。この点については次節

*) 本シリーズ I 『統計学入門』第 9 章(9.5)参照。

の標本の大きさの決定のところで再びふれる。

母平均 μ に関する 2 標本問題* a) 分散が共通で既知の場合。母分散の等しいふたつの正規母集団について母平均を比較する問題を考えよう。いま X_1, \dots, X_m を $N(\mu_1, \sigma^2)$ に従う互いに独立な確率変数とし、 Y_1, \dots, Y_n を $N(\mu_2, \sigma^2)$ に従う互いに独立な確率変数とする。ここで $\sigma^2 = \sigma_0^2$ は既知で、また 2 つの母集団に共通であるとする。片側検定問題は $H_0 : \mu_1 \leq \mu_2$, $H_1 : \mu_1 > \mu_2$ であり、両側検定問題は $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$ である。この場合の検定統計量は

$$(6.15) \quad Z = \sqrt{\frac{mn}{m+n}} \cdot \frac{\bar{X} - \bar{Y}}{\sigma_0}$$

である。また片側検定および両側検定問題における有意水準 α の棄却域はそれぞれ(6.9)および(6.11)で与えられる。1 標本問題と同様に計算すると片側検定および両側検定の検出力関数は、 $\tau = \sqrt{mn/(m+n)}(\mu_1 - \mu_2)/\sigma_0$ として、それぞれ

$$(6.16) \quad \begin{aligned} \beta_\delta(\mu_1, \mu_2) &= 1 - \Phi(z_\alpha - \tau) \\ \beta_\delta(\mu_1, \mu_2) &= 1 - \Phi(z_{\alpha/2} - \tau) + \Phi(-z_{\alpha/2} - \tau) \end{aligned}$$

で与えられることがわかる。

b) 分散は共通だが未知の場合は、1 標本の場合と同じく、この場合は t 検定を行うことになる。合併分散 s^2 を

$$(6.17) \quad s^2 = \frac{1}{m+n-2} \left(\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2 \right)$$

とおくとき、2 標本 t 統計量は

$$(6.18) \quad t = \sqrt{\frac{mn}{m+n}} \cdot \frac{\bar{X} - \bar{Y}}{s}$$

で与えられる。そして、片側検定および両側検定の有意水準 α の棄却域はそれぞれ

$$(6.19) \quad t > t_\alpha(m+n-2), \quad |t| > t_{\alpha/2}(m+n-2)$$

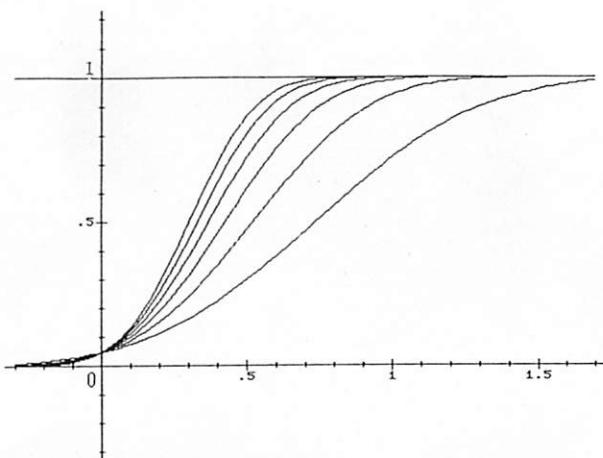
で与えられる。1 標本問題と同様、これらの検定の検出力関数はやや複雑であ

*) 以下の記述は羅列的であるので、順序を追って読む必要はなく、必要に応じて参照すればよい。

るが、近似的には σ^2 が既知の場合の(6.16)に等しい。

σ^2 に関する検定 μ が未知の場合、片側検定問題および両側検定問題は $H_0 : \sigma^2 \leq \sigma_0^2$, $H_1 : \sigma^2 > \sigma_0^2$ および $H_0 : \sigma^2 = \sigma_0^2$, $H_1 : \sigma^2 \neq \sigma_0^2$ である。平均からの偏差の二乗和を σ_0^2 で割ったものを

❖パソコン画面で見る検出力のようす❖



正規分布 $N(\mu, 1)$ についての、帰無仮説 $\mu=0$ の片側検定(有意水準 5%)の検出力
標本の大きさは、 $n=5(5)30$ である*).

対立仮説に属する θ (母数)ごとに検出力 $\beta(\theta)$ を計算し、これを曲線として表すと、検出力の曲線が得られる。これを、BASIC のプログラムによりパソコンで行い、画面に出すと、そのようすが一目でわかる。標本の大きさ n を自由に変化させると、その曲線が何本も描けて、検出力の意味もよくわかり、また何よりも次節で扱う適切な標本の大きさ n を選ぶのに容易となる。

とはいって、検出力の計算には、正規分布 Φ やそれから導かれた諸分布(χ^2 , t , F)の確率計算の計算公式、画面出力のソフトウェアなどが必要である。これには、数学的な数値計算法の知識がかなり要求されるので、誰にでもすぐにできるというわけではない。画面をかざる検出力の何本ものきれいな曲線が誰にでも手軽に描ける、というのはまだのようである。

*) 数表・図表の表記法で、間の間隔 5 で 5~30 の意。

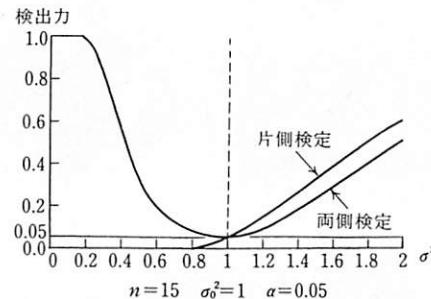


図 6.4 正規分布の(母)分散の検定の検出力
母平均についての同様の検出力図 6.3 と類似の形をとっている。やはり帰無仮説 $\sigma^2=1$ のところで有意水準の値 0.05 に一致している。この他、両側検定、片側検定のそれぞれについて、図 6.3と同じことが言える。

$$(6.20) \quad \chi^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2}$$

とおく。片側検定および両側検定の有意水準 α の棄却域はそれぞれ

$$(6.21) \quad \begin{aligned} & \chi^2 > \chi_{\alpha}^2(n-1) \\ & \chi^2 > \chi_{\alpha/2}^2(n-1) \text{ または } \chi^2 < \chi_{1-\alpha/2}^2(n-1) \end{aligned}$$

で与えられる。ただし $\chi_{\alpha}^2(\nu)$ は自由度 ν の χ^2 分布 $\chi^2(\nu)$ の上側 α 点を表す。ここで片側検定の検出力関数を具体的に計算してみると次のようにになる。

$$(6.22) \quad \begin{aligned} \beta_{\alpha}(\sigma^2) &= P_{\sigma^2}(\chi^2 > \chi_{\alpha}^2(n-1)) \\ &= P_{\sigma^2}\left(\frac{\sum(X_i - \bar{X})^2}{\sigma^2} > \frac{\sigma_0^2}{\sigma^2} \chi_{\alpha}^2(n-1)\right) \\ &= 1 - G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2}, \chi_{\alpha}^2(n-1)\right) \end{aligned}$$

となる。ただし、 G_{n-1} は自由度 $n-1$ の χ^2 分布 $\chi^2(n-1)$ の累積分布関数である。同様に両側検定の検出力関数は

$$(6.23) \quad \beta_{\alpha}(\sigma^2) = 1 - G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2}, \chi_{\alpha/2}^2(n-1)\right) + G_{n-1}\left(\frac{\sigma_0^2}{\sigma^2}, \chi_{1-\alpha/2}^2(n-1)\right)$$

で与えられる。

分散に関する2標本問題 μ_1, μ_2 がともに未知の場合、 X_1, \dots, X_m を $N(\mu_1, \sigma_1^2)$ からの標本、 Y_1, \dots, Y_n を $N(\mu_2, \sigma_2^2)$ からの標本とする。片側検定問題および両側検定問題はそれぞれ $H_0: \sigma_1^2 \leq \sigma_2^2$, $H_1: \sigma_1^2 > \sigma_2^2$ および $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$ と表される。用いる検定統計量 F は

$$(6.24) \quad F = \frac{\sum_{i=1}^m (X_i - \bar{X})^2 / (m-1)}{\sum_{i=1}^n (Y_i - \bar{Y})^2 / (n-1)}$$

である。片側検定および両側検定の有意水準 α の棄却域はそれぞれ

$$(6.25) \quad \begin{aligned} & F > F_{\alpha}(m-1, n-1) \\ & F > F_{\alpha/2}(m-1, n-1) \text{ または } F < F_{1-\alpha/2}(m-1, n-1) \end{aligned}$$

と表される。ただし、 $F_{\alpha}(\nu_1, \nu_2)$ は自由度 ν_1, ν_2 の F 分布の上側 α 点である。これらの検定の検出力関数はそれぞれ

$$(6.26) \quad \begin{aligned} \beta_{\alpha}(\sigma_1^2, \sigma_2^2) &= 1 - G_{m-1, n-1}\left(\frac{\sigma_2^2}{\sigma_1^2} \cdot F_{\alpha}(m-1, n-1)\right) \\ & \beta_{\alpha}(\sigma_1^2, \sigma_2^2) = 1 - G_{m-1, n-1}\left(\frac{\sigma_2^2}{\sigma_1^2} \cdot F_{\alpha/2}(m-1, n-1)\right) \\ & \quad + G_{m-1, n-1}\left(\frac{\sigma_2^2}{\sigma_1^2} \cdot F_{1-\alpha/2}(m-1, n-1)\right) \end{aligned}$$

で与えられる。ただし G_{ν_1, ν_2} は自由度 ν_1, ν_2 の F 分布 $F(\nu_1, \nu_2)$ の累積分布関数である。

6.2.2 二項分布の成功確率に関する検定

X を二項分布 $Bi(n, p)$ に従う確率変数とする。 p は成功確率で、関心のある未知母数である。片側検定問題および両側検定問題は

$$(6.27) \quad \begin{aligned} H_0: p &\leq p_0, & H_1: p > p_0 \\ H_0: p &= p_0, & H_1: p \neq p_0 \end{aligned}$$

である。用いる検定統計量は

$$(6.28) \quad Z = \sqrt{n} \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}}, \quad \hat{p} = \frac{X}{n}$$

である。片側検定および両側検定の近似的に有意水準 α の棄却域は (6.11) および (6.9) とすればよい。この場合、 Z を標準正規分布のパーセント点と比較するのは中心極限定理によって (6.28) の Z が漸近的(近似的)に標準正規分布に従うことによる。中心極限定理による結果は n が大きく、また p が 0 または 1 にあまり近くなければ正当化される。ただし、これはあくまで近似的な結果であり有意水準も近似的に α であるという点に注意しなければならない。

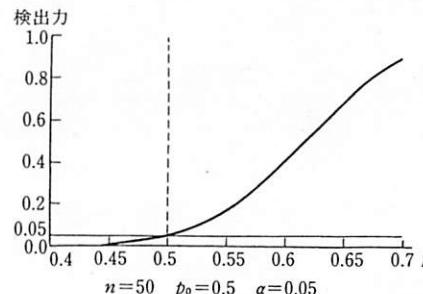


図 6.5 二項分布の成功確率 p の検定の検出力
(6.28)による二項分布の成功確率に関する検定は、簡単な場合の教育例として、図 5.1 に検出力が示されているが、これは一般的な場合の正規近似(6.29)をした結果の検出力である。(6.12)(6.29)を比べれば、本図が図 6.3 と非常に類似していることは理解されよう。

ここで片側検定の検出力を求めてみよう。ただし、これも中心極限定理を用いた近似的な結果である。

$$(6.29) \quad \beta_\delta(p) = P_p\left(\frac{\sqrt{n}(\hat{p}-p_0)}{\sqrt{p_0(1-p_0)}} > z_\alpha\right)$$

$$= P_p\left(\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}} > \frac{z_\alpha\sqrt{p_0(1-p_0)} - \sqrt{n}(p-p_0)}{\sqrt{p(1-p)}}\right)$$

$$\doteq 1 - \Phi\left(\frac{z_\alpha\sqrt{p_0(1-p_0)} - \sqrt{n}(p-p_0)}{\sqrt{p(1-p)}}\right)$$

となる。両側検定の検出力関数は省略する。

2 標本問題 X を二項分布 $Bi(m, p_1)$ に従う確率変数、 Y を二項分布 $Bi(n, p_2)$ に従う確率変数とし、これらは独立とする。片側検定問題および両側検定問題は、 $H_0 : p_1 \leq p_2$, $H_1 : p_1 > p_2$ および $H_0 : p_1 = p_2$, $H_1 : p_1 \neq p_2$ である。いま $\hat{p}_1 = X/m$, $\hat{p}_2 = Y/n$, $\hat{p} = (X+Y)/(m+n)$ とおき、検定統計量 Z を

$$(6.30) \quad Z = \sqrt{\frac{mn}{m+n}} \cdot \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})}}$$

とおく。片側検定問題および両側検定問題の近似的に有意水準 α の棄却域は、この場合も(6.11)および(6.9)で与えられる。検出力関数は複雑となるが、 p_1 と p_2 があまり離れていないという前提のもとでラフな近似を行うと、片側検定の検出力は

$$(6.31) \quad \beta(p_1, p_2) = 1 - \Phi\left(z_\alpha - \frac{p_1 - p_2}{\sqrt{p_1(1-p_1)/m + p_2(1-p_2)/n}}\right)$$

で与えられる。

6.3 標本の大きさ n の決定

検出力関数の重要な応用は、必要な検出力を確保するための標本の大きさ n をどのように決めるかということである。

例 6.4 〈不良率の検定に必要な n 〉 不良率に関する例 6.1 をとりあげよう。いま検定方式 δ_2 を考えよう。まず、すでに表 6.1 で見たように、この検定は有意水準 5% の検定((6.4) 参照)である。つぎに、表 6.1 より δ_2 の $p=0.03$ での検出力は 0.353 である。すなわち $p=0.03$ であって製造工程が異常であっても、 δ_2 を用いてそのような異常を見逃す過誤(すなわち第 2 種の過誤)の確率は $1-0.353=0.647$ とかなり大きい。もし $p=0.03$ という不良率がかなり深刻な状況を表しており、このような事態をより大きな確率で検出したいと考えるならば、 n を大きくしなければならない。

いずれ示すように $n=100$ と固定すれば、有意水準を変更することなしに $p=0.03$ における検出力を高くすることは不可能だからである。いま $n=200$ とし検定方式として今度は 5 個以上の不良品があった場合に帰無仮説を棄却する検定方式 δ_3 を考えよう。この検定について検出力関数を計算してみると、第 1 種の過誤の最大値が $\beta_{\delta_3}(0.01) = 0.0517$ とほぼ有意水準 5% であり、また $\beta_{\delta_3}(0.03) = 0.719$ となり $p=0.03$ における第 2 種の過誤の確率が $1-0.719=0.281$ と δ_2 に比べてかなり小さくなっていることがわかる。

このように標本の大きさ n を大きくすることによって対立仮説のもとで検出力を高くすることができる。ここでもう一つ次のような例を考えてみよう。

例 6.5 〈薬効検定〉 いま標準薬を用いたときの有効率を p_0 とし、新標準薬を用いたときの有効率を p とする。たとえば $p_0=0.5$ であることは、薬が長く使われてきたことから既知であるとしよう。また $p=0.6$ 程度であると期待されているとする。この場合 $p > p_0$ であるから、新薬は標準薬より有効と期待されているのであるが、このことを実験データによって検証するには次のような検定を行う必要がある。

すなわち $H_0 : p \leq p_0$ を帰無仮説とする検定を行って、帰無仮説が棄却できればよい。さて、 n が小さいと、かりに $p=0.6$ であっても検出力が小さいために帰無仮説が棄却できない可能性が大きい。直観的にも、 $p=0.6$ と $p_0=0.5$

の差が10%とそれほど大きくないために、標本の大きさ n の小さい実験では、この有効率の差を十分に証明できないと思われる。そこである程度 n の大きい実験を行う必要がある。前節の二項分布の成功確率に関する片側検定の検出力関数の結果を用いると、必要な n を具体的に求めることができる。

いま $p=0.6$ が正しいときに、検出力を0.8程度にしたいとしよう。(6.29)において $p_0=0.5$, $p=0.6$, $\alpha=0.05$ として、検出力=0.8という式

$$\begin{aligned}\beta_\alpha(0.6) &= 1 - \Phi\left(\frac{z_\alpha \sqrt{p_0(1-p_0)} - \sqrt{n}(p-p_0)}{\sqrt{p(1-p)}}\right) \\ &= 1 - \Phi\left(\frac{1.64/2 - \sqrt{n} \cdot 0.1}{\sqrt{0.6 \times 0.4}}\right) \\ &= 0.8\end{aligned}$$

を解けばよい。これから

$$z_{0.8} = -0.84 = \frac{0.82 - \sqrt{n} \cdot 0.1}{\sqrt{0.24}}$$

となる。これを解くと n は、少くとも

$$(6.32) \quad n = 152$$

となる。標本の大きさ n は実験をする前に決定しなければならないので、このような n の計算は応用上も重要であることが理解できるだろう。

このように検出力関数を n に関して解くことによって必要な n を求めることができるが、局外母数が含まれる場合には、必要な n の決定はもう少しやっかいである。局外母数 *nuisance parameter*^{*)} とは興味のある母数以外の未知母数を言う。たとえば、正規分布の母平均 μ に関する検定を行いたいとき σ^2 が未知であるとする。このとき σ^2 は局外母数である。

検出力関数は通常局外母数にも依存する。これが‘やっかいさ’の原因である。したがって n を決定するには、局外母数に関する事前の推定値がなければならない。たとえば正規分布の μ に関する検定においては、 n を決定するには σ^2 の事前推定値 $\hat{\sigma}^2$ が必要である。このような事前の推定値が得られない場合には、まず予備的な小規模の実験を行い σ^2 の推定値を得てから、本格的な実験の標本の大きさを決定するという2段階の手続きが必要である。

*) *nuisance* は、「妨げ」「迷惑」「邪魔」の意味。

いろいろな検定の n 前節でとりあげた主な検定について、検出力関数を n について解くことにより必要な n を求めよう。ここでは、スペースの関係上、応用上最も重要な2つの場合について結果のみを示すこととする。また、結果は片側検定の場合について示す。両側検定の場合には有意水準を2で割って有意水準 $\alpha/2$ の片側検定の結果を用いればよい。たとえば有意水準5%の両側検定で必要な n を求めなければ、有意水準2.5%の片側検定を考えて必要な標本を求めればよい。

a) μ に関する片側検定 (σ^2 は未知) のとき、 $\hat{\sigma}_0^2$ を σ^2 に対する事前の推定値とする。いま特定の対立仮説 $\mu > \mu_0$ のところで、 γ ($0 < \gamma < 1$) 以上の検出力を確保したければ、標本の大きさを

$$(6.33) \quad n \geq \frac{\hat{\sigma}_0^2(z_\alpha + z_\gamma)^2}{(\mu - \mu_0)^2}$$

とすればよい。この式は時おり用いられる。

b) 二項分布の成功確率 p に関する片側検定のとき、 $p > p_0$ となる対立仮説の点で、同じく γ 以上の検出力を確保したいとする。これはすでに上の例で見た場合である。 n の一般式は

$$(6.34) \quad n \geq \frac{(z_\alpha \sqrt{p_0(1-p_0)} - z_\gamma \sqrt{p(1-p)})^2}{(p - p_0)^2}$$

とすればよい。

例 6.6 〈続〉 例 6.5 で論じた薬効検定の場合の標本の大きさを確認してみよう。 $p_0=0.5$, $z_{0.05}=1.64$, $p=0.6$, $\gamma=0.8$, $z_\gamma=z_{0.8}=-0.84$ を代入すれば、ただちに

$$n \geq 100(1.64/2 + 0.84\sqrt{0.24})^2 = 151.7$$

となり、(6.32)が確認できる。

6.4 最強力検定

すでに1節で説明したように検定の良さを比較するには、第1種の過誤の確率を最小にした上で、対立仮説のもとでの検出力を最大にすればよい。最強力検定の理論は、このような意味で最も望ましい検定方式をどのように構成する

かということに関する理論である。この理論により t 検定などの標準的な検定方法の最適性が保証される。

一般の検定問題を $H_0: \theta \in \Theta_0, H_1: \theta \in \Theta_1$ とする。有意水準 α の

すべての $\theta \in \Theta_0$ に対し $\beta_\delta(\theta) \leq \alpha$,

の検定の中で、特定の対立仮説の点 $\theta_1 \in \Theta_1$ において検出力を最大

$$\max_{\delta} \beta_\delta(\theta_1) = \beta_{\delta^*}(\theta_1)$$

にする検定 $\delta_{\theta_1}^*$ を、 θ_1 における最強力検定 most powerful test という。また、ある検定方式 δ^* がいかなる対立仮説の点 $\theta_1 \in \Theta_1$ においても同時に最強力検定であるとき δ^* を一様最強力検定 uniformly most powerful test という。一様最強力検定が存在すればそれが最も望ましい検定である。

ところで、帰無仮説および対立仮説がともに 1 点である場合が最も単純な検定問題である。1 点、すなわち単独の分布からなる仮説を単純仮説 simple hypothesis とよぶ。実は単純帰無仮説を単純対立仮説に対して検定する検定問題においては最強力検定を明示的に構成できる。この構成法を与えるのが有名なネイマン・ピアソンの補題 Neyman-Pearson's lemma である^{*)}。密度関数あるいは確率関数を $f_\theta(x)$ とする。記法の簡単のために、ここでは補題を $n=1$ 、すなわち $f_\theta(x)$ に従う確率変数 X が 1 個得られるとした場合について述べているが、一般的 n に対しても $f_\theta(x)$ を $\prod_{i=1}^n f_\theta(x_i)$ でおきかえれば全く同じ結果が成り立つ。

〈ネイマン・ピアソンの補題〉 単純帰無仮説を単純対立仮説に対して検定する検定問題

$$(6.35) \quad H_0: \theta = \theta_0, \quad H_1: \theta = \theta_1$$

を考える。いま $k \geq 0$ に対して

$$(6.36) \quad \begin{aligned} \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} > k \text{ ならば帰無仮説を棄却,} \\ \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} \leq k \text{ ならば帰無仮説を採択} \end{aligned}$$

で定義される検定方式を δ^* とおく。いま、定数 k を δ^* の第 1 種の過誤の確

^{*)} Pearson は Karl Pearson の子 Egon S. Pearson.

率が、指定された α に対して

$$(6.37) \quad \alpha = P_{\theta_0}(f_{\theta_1}(X)/f_{\theta_0}(X) > k)$$

となるようにおこう。このとき、有意水準 α の検定の中で δ^* は θ_1 における検出力を最大にする。すなわち δ^* は最強力検定である。

証明 ここでは $f_\theta(x)$ が密度関数の場合について証明する。 $f_\theta(x)$ が確率関数の場合も同様である。証明中、表記を簡単にするために $f_\theta(x) \equiv f(x, \theta)$ と記そう。

いま δ を任意の有意水準 α の検定方式とし B を δ の棄却域とする。このとき

$$\int_B f(x, \theta_0) dx = \beta_\delta(\theta_0) \leq \alpha$$

である。したがって、

$$k \left(\alpha - \int_B f(x, \theta_0) dx \right) = k \left(\int_{B^*} f(x, \theta_0) dx - \int_B f(x, \theta_0) dx \right) \geq 0$$

となる。ただし B^* は δ^* の棄却域である。これにより

$$(6.38) \quad \begin{aligned} & \beta_{\delta^*}(\theta_1) - \beta_\delta(\theta_1) \\ &= \int_{B^*} f(x, \theta_1) dx - \int_B f(x, \theta_1) dx \\ &\geq \int_{B^*} f(x, \theta_1) dx - \int_B f(x, \theta_1) dx - k \left(\int_{B^*} f(x, \theta_0) dx - \int_B f(x, \theta_0) dx \right) \\ &= \int_{B^*} (f(x, \theta_1) - kf(x, \theta_0)) dx - \int_B (f(x, \theta_1) - kf(x, \theta_0)) dx \\ &= \int_{B^* \cap B^c} (f(x, \theta_1) - kf(x, \theta_0)) dx - \int_{B \cap (B^*)^c} (f(x, \theta_1) - kf(x, \theta_0)) dx \end{aligned}$$

となる。ただし A^c は A の補集合を表し \cap は共通集合を表す。ところで B^* の定義から $B^* \cap B^c$ においては $f(x, \theta_1) - kf(x, \theta_0) > 0$ であり、また $B \cap (B^*)^c$ においては $f(x, \theta_1) - kf(x, \theta_0) \leq 0$ となる。したがって、(6.38) の右辺は非負である。これより $\beta_{\delta^*}(\theta_1) \geq \beta_\delta(\theta_1)$ となるから、 δ^* は最強力検定となる。(終)

例 6.7 〈二項分布に対する一様最強力検定〉 補題を二項分布 $Bi(n, p)$ に応用してみよう。 $p_0 < p_1$ とし、 p_0 対 p_1 の検定を考える。このときネイマン・ピアソンの補題より最強力検定の棄却域は

❖ショッピングをしながら統計学の大定理を学ぶ。

「ネイマン・ピアソンの補題」は、最強力検定、一様最強力検定を作るための補助定理という意味で「補題」とよばれているが、実はこれを補題というのではなくていい。『定理』とか『大定理』とよばれてもおかしくない。人生(?)や現実世界の基本原理を表しているのである。

——一定額の限られた予算の制約の下で、お金の最大限の利用を引き出そうと望んでいる買手は、品物をドル当たりの価値 value per dollar の順序にランク付けすることからはじめるだろう。与えられた距離を可能な限り短時間で行こうとするなら、交通手段の中で最も速い方法を使うであろう。つまり、時間当たりのマイル数 miles per hour で見て、最大の数を与えるものを選ぶであろう。すなわち、最も価値のある点 x とは、一般に比

$$r(x) = f_{\theta_1}(x)/f_{\theta_0}(x)$$

が最大になる点である。

この比によって各点をランキングして上位から選び出して行き、ちょうど、制約条件 α いっぱいになるまで(そこを k とする)統ければよい。——

レーマン E. L. Lehmann というアメリカの統計学者は、基本補題 fundamental lemma とよんだうえ、このようにわかり易く説明している。

デパートでショッピングするとき、人生の総価値を人間の時間制約 75 年の下で考えようとするとき、この補題が役に立ちそうである。

$$(6.39) \quad \frac{\binom{n}{x} p_1^x (1-p_1)^{n-x}}{\binom{n}{x} p_0^x (1-p_0)^{n-x}} = \left(\frac{1-p_1}{1-p_0}\right)^n \left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)^x > k$$

で与えられる。 $p_1 > p_0$ より $p_1(1-p_0)/[p_0(1-p_1)] > 1$ であるから、対数をとればわかるように、(6.39)の不等式は、ある(k から決まる)定数 c を用いた

$$(6.40) \quad x > c$$

の形の不等式と同値である。したがって、(6.40)を棄却域とする検定が最強力検定となる。(6.40)は当たり前のように見えるが、ここの主張はこの検定方式が一番よい検定(最強力検定)である、ということである。

ところで c は有意水準が与えられた確率(たとえば 0.05)となるように決めさえすればよいから、実は対立仮説に依存せずに決まっている。すなわちどの対立仮説を考えても(6.40)の検定が最強力検定になっていることがわかる。し

たがって、二項分布における片側検定 $H_0: p \leq p_0, H_1: p > p_0$ を考えれば(6.40)の検定が一様最強力検定となっていることがわかる。(このことについては、(6.1)で不良品に関する例に関してすでに述べておいた。)

正規分布の場合も、母分散が既知の場合の母平均に関する片側検定および母平均が既知の場合の母分散に関する片側検定は一様最強力検定であることが証明される。このように正規分布や二項分布における片側検定では通常の検定が一様最強力検定となる。

不偏検定 しかしながら、両側検定問題や局外母数が存在する場合には一様最強力検定が存在しないことが多い。それは対立仮説が帰無仮説のどちらに来るかによって、最強力検定の形がかわってしまうからである。このような場合には考慮する検定を不偏性を持つ検定に限れば、一様最強力検定が存在することがある。このような検定を一様最強力不偏検定という。

有意水準 α の検定 δ が不偏 unbiased であるとは、

$$\text{任意の } \theta \in \Theta_1 \text{ に対し, } \beta_\delta(\theta) \geq \alpha$$

が成り立つことをいう。対立仮説のもとで検出力を大きくするのがよい検定であるから、任意の対立仮説のもとで検出力が有意水準以上であることを要求するのは自然な要求であると思われる。

いま不偏検定に限って議論すれば、たとえば正規分布に関する t 検定は一様最強力不偏検定であることが示される。二項分布の成功確率に関する両側検定、正規分布の母分散に関する両側検定、二項分布の成功確率に関する検定などは近似的に一様最強力不偏検定であることが示される。(図 6.3, 6.4 参照)

◆◆◆◆ 第6章 練習問題 ◆◆◆◆

6.1 <二項分布に対する検定の検出力> コインにゆがみがないかどうかを検定するのに、そのコインを 4 回投げて 4 回とも表が出るか 4 回とも裏が出たらゆがみがあるとする検定方式を考える。この検定方式の検出力関数を求めグラフにせよ。

6.2 <母平均についての 2 標本問題> 正規分布の母平均 μ に関する 2 標本問題で、(6.16)を用いて、 $m=n=5, \mu_1-\mu_2=1, \sigma_0=1$ の場合の検出力を求めよ。

6.3 <同、証明> 正規分布の母平均に関する 2 標本問題において、(6.15)の検

定の検出力は(6.16)で与えられることを示せ。

6.4 <二項確率の検定の標本の大きさ> 2項分布の成功確率の片側検定で, $\alpha=0.4$ とし, $\beta=0.6$ のときの検出力を 0.9 以上にしたい。このために必要な標本の大きさを求めよ。

6.5 <正規分布の母平均の一様最強力検定> ネイマン・ピアソンの補題を用いて, 正規分布で $\sigma^2 = \sigma_0^2$ が既知の場合の μ に関する片側検定で, (6.11)の検定が一様最強力検定となることを証明せよ。

第7章 分布の仮定

これまでに説明した推定や検定の方法は, データが正規分布をはじめとする一定の確率分布に従うという仮定のもとで導き出された。しかし, 正規分布の仮定が満たされないときもある。そこで, 正規分布を前提とした推定や検定の方法が受ける影響を調べ, 分布の仮定に基づかない, ノンパラメトリック法とよばれる推測の方法を紹介し, あわせて正規分布の仮定が正しいかどうかをチェックする方法を説明する。

7.1 正規分布の仮定

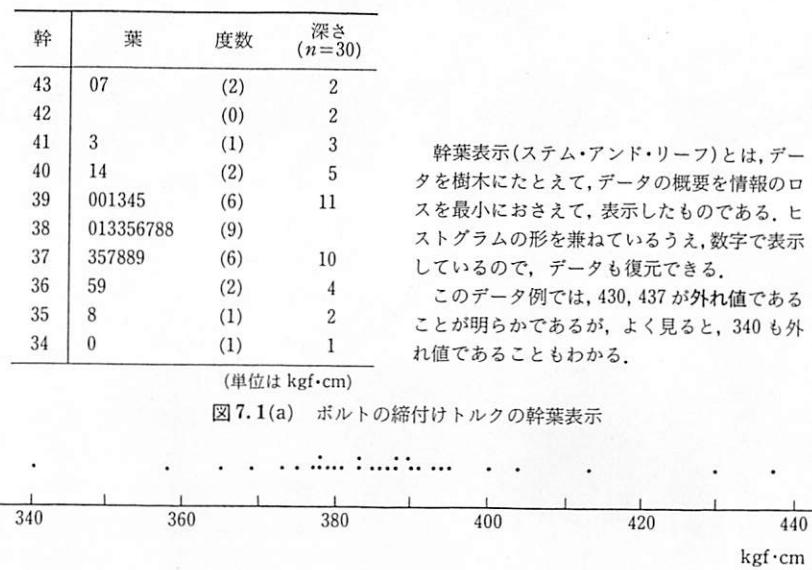
7.1.1 はじめに

これまでの章で述べた様々な統計的推測の方法, すなわち推定や検定の方法は, 得られるデータが一定の確率分布に従うという仮定の下で導き出されたものであった。とくに, データが連続量の場合には, その分布が正規分布であると仮定されることが多い。しかし, 現実のデータは必ずしもこのような仮定をみたすとは限らない。次の 2 つのデータは, 正規分布からのものであることが疑わしい例である。

例 7.1[†] <締め付けトルク> 下記のデータはある工業製品の組立工程で, 締付け機で締め付けたボルトの締付けトルク (kgf·cm) を $n=30$ 個について測定したものである。

395	383	365	381	437	378	394	401	385	379
413	387	369	391	340	375	393	390	390	380
388	388	383	386	404	430	377	373	378	358

図 7.1(a), (b) は, このデータを幹葉表示 stem and leaf display および数直



線表示によって表したものである。図 7.1(a)では、たとえば、5番目の437という数値は、主要な桁(幹)である43と、その次の1桁(葉)である7に分けて表されている。その右側の()の中は各幹の度数である。右端の「深さ」は最大あるいは最小からの累積を示し、中央値や四分位点を求めるのに利用される。図 7.1(a)からはもとの数値が完全に復元できる。幹葉表示の目的は、ヒストグラムと同じくデータ全体の状況を把握することであるが、ヒストグラムと違ってもとのデータのもつ情報がほぼ完全に保存されるという特長をもつ。

これらの図をみると、最大値437とその次に大きい430が他の値から飛び離れていることがわかる。このように、他のデータから飛び離れたデータは外れ値 outlier とよばれる。最小値340も他から飛び離れている。このデータでは、全体のうちかなりの割合が中央部に集中する反面、それから大きく離れた数値も両側に存在している。それは正規分布の確率密度関数から想像されるものとは異なる。

もちろん、このようなデータが正規分布から得られる可能性もないわけではない。第11章で乱数の知識を学んで、パソコンでシミュレーションをしてもらえばすぐにわかるように、とくにデータ数が小さいときは、正規分布からの

データが「正規分布からのものらしく」みえる方がむしろ稀であるといってよい。しかしながら、上記の特徴はデータが正規分布からのものであることを疑わせるものである。なお、データが正規分布からのものであることをチェックする方法は7.4節で説明する。

例 7.2 <ラットの肺重量> 下記のデータは約2週齢のラット19匹の肺の重量(g)を測定したものである。

0.378	0.421	0.324	0.388	0.333	0.384	0.334	0.359	0.369	0.360
0.384	0.374	0.363	0.484	0.321	0.322	0.329	0.343	0.352	

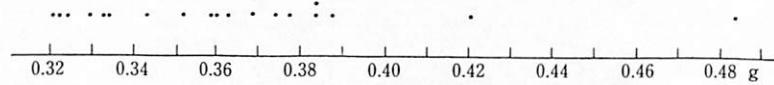
(出典：水内知子・木田厚瑞氏(東京都老人医療センター)の御好意による。)

このデータの幹葉表示および数直線表示を図 7.2(a), (b)に示した。この場合、数値421が他からかなり飛び離れているが、その上にさらに484という外れ値があり、「大きい」という幹に()に入れて表されている。このように、データが右にすそを引くような非対称性を示すことは、生体に関する測定値ではむしろありふれたことである。

幹	葉	度数	深さ (n=19)
大きい	(484)	(1)	1
42	1	(1)	2
41		(0)	2
40		(0)	2
39		(0)	2
38	448	(3)	5
37	48	(2)	7
36	039	(3)	
35	29	(2)	9
34	3	(1)	7
33	34	(2)	6
32	1249	(4)	4

(単位は 0.001 g)

図 7.2(a) ラットの肺重量の幹葉表示



正規性の仮定の非成立を見ることがある。

さて、データの真の分布が仮定した分布と異なっているとき、その仮定を前提にして導かれた検定や推定の方式は必ずしも妥当でなかったり、あるいは効率が悪くなったりするかもしれない。たとえば、7.2節で述べるように、母平均の推定に通常用いられる標本平均は、正規分布のもとでは他の推定量より分散が小さいが、真の分布が正規分布でないときは他の不偏推定量よりずっと分散が大きくなることがある。

そこで、真の分布が正規分布でないとき、正規分布を前提とした方法が受けける影響について調べてみよう。

7.1.2 繰り返し測定のモデル

いま、ある量を繰り返し測定するものとする。 n 回の測定から得られる観測値を X_1, X_2, \dots, X_n とし、それらが真の値^{*} θ と測定誤差 ε_i の和として

$$(7.1) \quad X_i = \theta + \varepsilon_i, \quad i=1, 2, \dots, n$$

と表されるものとしよう。本章ではこの場合だけを扱う。

統計理論において、このような形で確率分布に関係している母数を位置母数 location parameter という。

この測定誤差について、次のことを仮定する。

1° $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ は互いに独立に同一の分布に従う確率変数である。

2° $E(\varepsilon_i) = 0, V(\varepsilon_i) = \sigma^2, i=1, 2, \dots, n$

3° $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ は正規分布に従う。

そのとき、次のことが成り立つ。

I. θ の点推定については、標本平均を $\bar{X} = \sum X_i / n$ とすると、推定量

$$(7.2) \quad \hat{\theta} = \bar{X}$$

は θ の最小分散不偏推定量 minimum variance unbiased estimator になる。すなわち、 θ や σ^2 の値が何であっても、 $\hat{\theta}$ は θ の他の任意の不偏推定量より分散が小さい。そして、その分散は $V(\hat{\theta}) = \sigma^2 / n$ である。

II. 仮説 $\theta = \theta_0$ を有意水準 α で検定するには、前章まででよく知られている

^{*}) 本章は、主に一般論を扱うので記号 θ を用いる。他章におけるように μ を用いることが多い。

✧オッカムの剃刀✧

「ひげ」を無意味、不必要的仮説や仮定にたとえ、無用のひげは正しい思考をさまたげる所以剃り落すべきというもの。別の言い方では思考の節減の原理 principle of parsimony, あるいは「節約の原理」「ケチの原理」とも言う。よく出てくる比喩であるが、統計理論に出てくる多くの仮定でも、どの仮定が本質的に必要かという自省は常に必要である。

少し固い話になるが、中世も終りに近い、イギリスのスコラ哲学家オッカムのウィリアム(William of Ockham, 1285-1349/50)は、「存在は必要もなく増してはならない」と述べた。もし、今日の「科学」というものに引き直して考えると、現象はそれを説明するために必要な最小限度の仮定や理論で説明されるべきもの、という主張となろう。それ以上のものは思考の正しさの妨げとなる。一つの実証的、科学的態度をあらわしたものである。

このように、オッカムは歴史が近世に入る前に近代的な科学的思考の原型をすでに持っていた人といわれる。ただし、彼が科学的思考を万能と考えたわけではない(実際、オッカムは神学者であった)。ついでながら、「かみそり」(razor)のたとえ自体はオッカムの創造ではない。

ように、

a) σ^2 が既知のときは

$$(7.3) \quad |\bar{X} - \theta_0| \geq z_{\alpha/2} \sigma / \sqrt{n}$$

のとき仮説を棄却する。ただし、 $z_{\alpha/2}$ は標準正規分布の上側 $100\alpha/2$ パーセント点である。

b) σ^2 が未知のときは、不偏分散を

$$(7.4) \quad s^2 = \sum (X_i - \bar{X})^2 / (n-1)$$

として

$$(7.5) \quad |\bar{X} - \theta_0| \geq t_{\alpha/2}(n-1) s / \sqrt{n}$$

のとき仮説を棄却する。ただし、 $t_{\alpha/2}(n-1)$ は自由度 $n-1$ の t 分布の上側 $100\alpha/2$ パーセント点である。

III. θ に対する区間推定として、信頼係数 $1-\alpha$ の信頼区間は、

a) σ^2 が既知のときは

$$(7.6) \quad \bar{X} - z_{\alpha/2} \sigma / \sqrt{n} < \theta < \bar{X} + z_{\alpha/2} \sigma / \sqrt{n}$$

b) σ^2 が未知のときは

$$(7.7) \quad \bar{X} - t_{\alpha/2}(n-1)s/\sqrt{n} < \theta < \bar{X} + t_{\alpha/2}(n-1)s/\sqrt{n}$$

となる。

以下では、仮定1°, 2°が正しいものとして、3°の正規分布の仮定が成り立たないとき、推測方式にどのような影響ができるかを、I, IIを中心に調べるとともに、分布の仮定にあまり影響されない推測方式について考えてみよう。

なお、正規分布の仮定をチェックする方法については、7.4節で述べよう。

7.2 点推定

7.2.1 すその長い分布

最初に点推定について考えよう。仮定1°, 2°が正しければ、明らかに

$$E(\bar{X}) = \sum E(X_i)/n = \theta, \quad V(\bar{X}) = \sum V(X_i)/n^2 = \sigma^2/n$$

であるから、正規性の仮定が成り立たなくとも標本平均 \bar{X} は θ の不偏推定量であり、その分散が σ^2/n あることに変わりはない。

しかし、 \bar{X} が最小分散になるとは限らない。実は、 $n \geq 3$ のとき、 \bar{X} が最小分散になるのは、 ε_i が正規分布に従う場合に限られ、 ε_i の分布が正規分布でないときは、分散が \bar{X} より小さい不偏推定量が存在することが理論的に証明されている。

誤差 ε_i が正規分布に従わない場合の一つの想定として、その分布が次の密度関数 f_t をもつと仮定しよう。

$$(7.8) \quad f_t(x) = \frac{2}{\pi} \cdot \frac{1}{(1+x^2)^{3/2}}, \quad -\infty < x < \infty$$

この分布は、自由度3の t 分布 $t(3)$ を分散が1になるように標準化したものである

$$E(\varepsilon_i) = 0, \quad V(\varepsilon_i) = 1, \quad i=1, 2, \dots, n$$

が成り立つ。

この分布を平均0、分散1の正規分布 $N(0, 1)$ の密度関数

$$(7.9) \quad \phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad -\infty < x < \infty$$

と比較したのが図7.3である。この図からは分布(7.8)の方が正規分布よりも

7.2 点推定

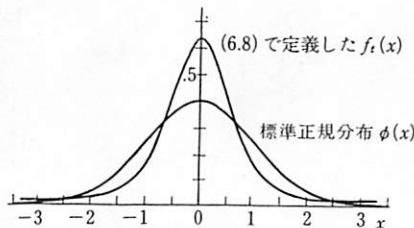


図7.3 検定誤差 ε_i の確率分布

(7.8)の $f_t(x)$ と標準正規分布 $\phi(x)$ 。両者とも左右対称で、確率分布のメディアンは0となっている。よく見ると、 $f_t(x)$ の場合の方が大きい誤差が起りやすく、「すそが長い」といわれる。実際、 $x=3$ の周辺で f_t と ϕ の上、下は逆転し、 f_t の方が0より遠い値をとりやすくなっている。(くわしくは、図7.4参照)

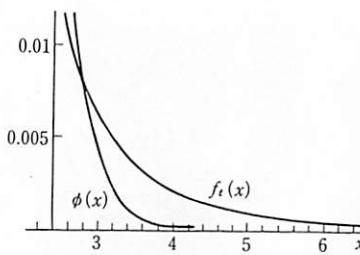


図7.4 x が大きいときの図7.3の拡大図

2つの誤差分布 f_t と ϕ のくわしい比較

図7.3では、最も重要な「すそ」が十分に見えないが、3の前後より先を拡大すると、 f_t の方がスソが長いことが一目瞭然となる。

らつきが小さいように見える。しかし、 x が大きい部分の拡大図7.4からわかるように、分布の両端のすその方では正規分布の方が密度関数の値がずっと速く0に近づく。実際、上側 $100\alpha/2$ (両側 100α) パーセント点、つまり、密度関数を一般に $f(x)$ として

$$\int_{-c_{\alpha/2}}^{c_{\alpha/2}} f(x) dx = 1 - \alpha$$

を満たす $c_{\alpha/2}$ を比較すると、表7.1のようになっている。すなわち、正規分布の場合には絶対値が3.29を越えるような値は1000回に1回しか起こらないのに対し、分布(7.8)の場合にはほぼ100回に1回起こることになり、また1000回に1回は絶対値が7.46を越えるような大きな誤差が起こる。

このような現象を、この分布は正規分布よりもすそが長い long-tailed と表現する。すそが長い分布の代表的なものは、コーシー分布(自由度1の t 分布

表7.1 正規分布と分布(7.8)の両側 100α パーセント点

α	0.001	0.01	0.05	0.10
正規分布	3.291	2.576	1.960	1.645
分布(7.8)	7.464	3.372	1.837	1.360

下段の確率分布の方が確率が遠方(すそ)まで残っている。

t(1))

$$(7.10) \quad f_t(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$$

である^{*)}が、この分布は原点について対称であるのに平均をもたない。すなわち

$$\int_{-\infty}^{\infty} xf(x) dx = \int_{-\infty}^{\infty} \frac{x dx}{\pi(1+x^2)} = \frac{1}{2\pi} [\log(1+x^2)]_{-\infty}^{\infty} = \infty - \infty$$

で、この値は不定となる。したがってまた、分散ももたないので、コーシー分布については仮定2°は成立しない。しかし、実際には、誤差が平均をもたないほどその長い分布に従うとは考えにくい。そこで、ここでは(7.8)をそのままの長い分布の代表として選んだわけである。

7.2.2 最尤推定量と標本中央値

さて、それが長い分布については、標本平均 \bar{X} は必ずしもよい推定量になるとは限らない。すなわち、 $\hat{\theta} = \bar{X}$ とすると

$$(7.11) \quad V(\hat{\theta}) = V(\bar{X}) = \sigma^2/n$$

であるが、これより分散の小さい θ の不偏推定量が存在する。分布(7.8)の f_t の場合を例にとって、最尤推定量と標本中央値の分散が標本平均より小さくなることを示してみよう。

最尤推定量 観測値 X_1, \dots, X_n が未知母数 θ を含む確率分布に従うものとして、その密度関数を表す。4.4節で説明したように、標本数 n が十分大きければ、 θ の最尤推定量 $\hat{\theta}_{ML}$ (第4章で $\tilde{\theta}$ と記したもの。MLとして最尤推定量であることを強調)はほぼ平均 θ 、分散 $1/nI(\theta)$ の正規分布に従う。ここで

$$(7.12) \quad I(\theta) = E_{\theta}\{\partial \log f_{\theta}(x)/\partial \theta\}^2 \\ = \int \{\partial \log f_{\theta}(x)/\partial \theta\}^2 f_{\theta}(x) dx$$

はフィッシャー情報量である。

さて、誤差 ε_i が分布(7.8)の f_t に従うとき、 X_i の分布は

$$(7.13) \quad f_{\theta}(x) = f_t(x-\theta) = \frac{2}{\pi(1+(x-\theta)^2)^2}$$

*) f_t と f_{θ} を混同しないこと。添字 t は t 分布にちなんで用いた関数名のためのものである。

となる。そのフィッシャー情報量は以下のように計算される。まず

$$(7.14) \quad \frac{\partial \log f_{\theta}(x)}{\partial \theta} = \frac{4(x-\theta)}{1+(x-\theta)^2}$$

となることから

$$I(\theta) = \int_{-\infty}^{\infty} \frac{16(x-\theta)^2}{(1+(x-\theta)^2)^2} \cdot \frac{2}{\pi(1+(x-\theta)^2)^2} dx$$

と表される。ここで、 $x-\theta$ を改めて x とおいて整理すると

$$(7.15) \quad I(\theta) = \frac{32}{\pi} \int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^4} dx \\ = \frac{32}{\pi} \left\{ \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^3} dx - \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^4} dx \right\} \\ = \frac{32}{\pi} \left(\frac{3\pi}{8} - \frac{5\pi}{16} \right) = 2$$

が得られる。

したがって n が大きいとき、 $V(\hat{\theta}_{ML}) \approx 1/(2n)$ となる。これは $V(\bar{X}) = 1/n$ のほぼ $1/2$ となっている。

最尤推定量の求め方 この場合の最尤推定量は次のようにして求められる。最尤推定量 $\hat{\theta}_{ML}$ の満たすべき方程式(尤度方程式)は

$$\frac{\partial}{\partial \theta} \left\{ \sum_i \log f_{\theta}(x_i) \right\} = 0,$$

と書かれるが、これは(7.14)より、

$$(7.16) \quad \sum_i \frac{x_i - \theta}{1 + (x_i - \theta)^2} = 0$$

に帰着する。(7.16)の解 θ は直接には得られないが、適当な近似解から出発して、逐次近似によって求めることができる。いま、

$$W_i = \frac{1}{1 + (x_i - \theta)^2}, \quad i=1, 2, \dots, n$$

とおくと、(7.16)は形式的に $\sum W_i(x_i - \theta) = 0$ 、したがって

$$\theta = \frac{\sum W_i x_i}{\sum W_i}$$

と表される。そこで、最初の近似解 $\theta^{(0)}$ を用いて

$$W_i^{(0)} = \frac{1}{1 + (x_i - \theta^{(0)})^2}, \quad i=1, 2, \dots, n$$

とおき

$$\theta^{(1)} = \frac{\sum W_i^{(0)} x_i}{\sum W_i^{(0)}}$$

を新しい近似解とする。このようにして、収束するまで次々と新しい近似解を計算すればよい。

標本中央値 観測値の分布が対称であれば、母平均 θ は母中央値 population median でもあるので、 θ の推定に標本中央値 sample median を用いることができる。そこで、観測値 X_1, X_2, \dots, X_n を大きさの順にならべた順序統計量 order statistics を

$$(7.17) \quad X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$$

であらわすとすると、標本中央値 X_{med} は

$$(7.18) \quad X_{\text{med}} = \begin{cases} X_{((n+1)/2)}, & n \text{ が奇数のとき} \\ (X_{(n/2)} + X_{((n/2)+1)})/2, & n \text{ が偶数のとき} \end{cases}$$

で与えられる。一般に標本中央値の n が大きいときの(すなわち漸近的)標本分布については、次の定理がよく知られている：

観測値 X_1, X_2, \dots, X_n の分布を $g(x)$ とするとき、 n が大きければ、 X_{med} は平均が母中央値 θ 、分散が $1/4n\{g(\theta)\}^2$ の正規分布

$$(7.19) \quad N\left(\theta, \frac{1}{4n\{g(\theta)\}^2}\right)$$

にほぼ従う。

いま、(7.1)を仮定したのであるから、誤差分布が(7.8)の $f_t(x)$ であれば、 $g(x) = f_t(x - \theta)$ であり、また、あきらかに母中央値は $x = \theta$ となる。したがって、標本中央値の分散は

$$(7.20) \quad V(X_{\text{med}}) = 1/4n\{f_t(0)\}^2 = \pi^2/16n = 0.617/n$$

となって、最尤推定量の $1/2n$ よりは大きいが、標本平均の $1/n$ よりも小さくなる。こうして、誤差分布が(7.8)の場合、標本平均 \bar{X} の分散は最尤推定量 $\hat{\theta}_{\text{ML}}$ や標本中央値 X_{med} に比べてかなり大きくなることがわかる。

7.2.3 ロバストな推定量

現実の問題においては、必ずしも誤差が正規分布に従うとは仮定できない。

その場合、分布の厳密な形がわかっていて、上記のように最尤法を適用するなどして標本平均 \bar{X} よりよい θ の推定量を求めることができる。しかし、誤差分布が正規分布でない特定の形であることを仮定できる場合は、ほとんどないといってよい。

普通に考えられるのは、誤差分布にはほぼ対称な一山型が仮定でき、正規分布よりはそぞが長い分布(7.8)のような場合である。このような場合は、上に示したように、 \bar{X} よりも標本中央値 X_{med} の方がよい推定量になる可能性がある。

一方、誤差が正規分布に従うときの X_{med} の分散は、上に述べたことから

$$(7.21) \quad V(X_{\text{med}}) = 1/4n\{\phi(0)\}^2 = \pi/2n = 1.57/n$$

となり、 \bar{X} のほぼ 1.57 倍となる。したがって、誤差分布が完全に正規分布になつていなくても、それに近いときは \bar{X} の方がよい推定量になるであろう。

実際、誤差分布が正規分布に近い形であれば、 \bar{X} の分散はその分布のもとで最もよい推定量の分散に比べてそれほど大きくならない。したがって、誤差分布が正規分布とそれほど離れていないと思われるならば、正規分布からの乖離は無視してもよい。

しかし、正規分布から大きく離れている場合には、 \bar{X} の分散は最もよい推定量の分散に比べて著しく大きくなる危険性がある。また、そのような場合は X_{med} の方が分散が小さくなる。誤差分布が正規分布と離れているか否かを標本からチェックする方法は 7.4 節で説明する。

誤差分布が(7.8)のように正規分布よりそぞが長い場合には、表 7.1 で検討したように、時々非常に絶対値の大きい誤差が生じることがある。そのとき、標本平均 \bar{X} はそのような値に影響されて、大きい誤差を生じることになる。これに対し、標本中央値 X_{med} はそのような影響を受けることがない。

そこで、正規分布のもとでも X_{med} より分散が小さく、しかも大きな誤差の影響を受けにくいうような推定量が望ましいことになる。それには、たとえば次のような形が考えられる。

トリム平均 順序統計量(7.17)の大きい方と小さい方からそれぞれ k 個ずつを除いて、残りの $(n-2k)$ 個の値を平均したもの

$$(7.22) \quad \hat{\theta}_{\text{trim}}(k) = \frac{1}{n-2k}(X_{(k+1)} + \cdots + X_{(n-k)})$$

をトリム平均 trimmed mean という。 k の値としては、 n が小さければ1か2くらい、 n がある程度大きければ n の5%か10%とするのが普通である。

トリム平均は、体操競技の採点などでは従来より使われている。主観的な判断が伴う採点などの場合、見まちがいによる大きな誤差や‘えこひいき’による偏りが起こりやすいであろうから、適切な評価を得るために両端の値を取り除くことは十分合理的であろう。

トリム平均は正規分布のもとでも \bar{X} より著しく分散が大きくなることがなく、正規分布から著しく離れた分布のもとでも分散があまり大きくならないことが知られている。すなわち、それはかなり広い範囲の分布の形に対して比較的よい性質をもつ。このことをトリム平均は分布の形に関してロバスト robust であるといふ。ロバストは“頑健”“びくともしない”の意味である。

フーバー型推定量 ロバストな推定量としては、もう少し複雑な次のような方式も考えられる。いま、 $[y]$ を y を越えない最大の整数として

$$(7.23) \quad s = ([n/2] + 1)/2$$

するとき

$$(7.24) \quad d = X_{(n-s+1)} - X_{(s)}$$

を四分位範囲 interquartile range とよぶ。ただし、 s に 0.5 の端数が出るときの $X_{(s)}$ は $X_{(s-0.5)}$ と $X_{(s+0.5)}$ の平均を意味するものとする。さらに

$$c = 0.5d$$

として

$$(7.25) \quad [X_{\text{med}} - c, \quad X_{\text{med}} + c]$$

の範囲にある観測値の個数を m 、その合計を S 、さらに

$$m_U = X_{\text{med}} + c \text{ より大きい観測値の個数}$$

$$m_L = X_{\text{med}} - c \text{ より小さい観測値の個数}$$

として

$$(7.26) \quad \hat{\theta}_H = \{S + c(m_U - m_L)\}/m$$

を θ の推定量とするものである。これをフーバー型推定量 Huber's estimator という*。

例 7.3 <外れ値のあるデータへの適用> 例 7.1 の締付けトルクのデータ

*) Huber (P. J.) は、現代の統計学者の名。

($n=30$) では、標本平均は

$$\bar{X} = \sum_{i=1}^{30} X_{(i)} = 11591/30 = 386.37$$

である。標本中央値 x_{med} は、図 7.1(a) から

$$X_{\text{med}} = (X_{(15)} + X_{(16)})/2 = (385 + 386)/2 = 385.5$$

となる。

$k=1$ のトリム平均は

$$\hat{\theta}_{\text{trim}}(1) = \sum_{i=2}^{29} X_{(i)} / 28 = 10814 / 28 = 386.21,$$

また、 $k=3$ の(10%)トリム平均は

$$\hat{\theta}_{\text{trim}}(3) = \sum_{i=4}^{27} X_{(i)} / 24 = 9248 / 24 = 385.33$$

となる。

フーバー型推定量は、 $s = ([30/2] + 1)/2 = 8$ から、四分位範囲は

$$d = X_{(23)} - X_{(8)} = 393 - 378 = 15$$

したがって、 $c = 0.5d = 7.5$ となり、(7.25)の範囲は

$$[X_{\text{med}} - c, X_{\text{med}} + c] = [385.5 - 7.5, 385.5 + 7.5] = [378, 393]$$

となる。この範囲にあるデータは $x_{(8)} = 378$ から $x_{(28)} = 393$ までの $m = 16$ 個で、その合計は $S = 6160$ 、さらに $m_U = 7$, $m_L = 7$ だから

$$\hat{\theta}_H = \{6160 + 7.5(7-7)\}/16 = 385.0$$

となる。

標本平均が大きい 2 つの外れ値の影響を受けて比較的大きな値をとっているのに比べ、 $k=3$ のトリム平均やフーバー型推定量の値はそれより 1 以上小さめになっている。(データの標準偏差は $s = 4.34$ であるから、この違いは無視できないものと考えられる。)

7.3 仮説検定

7.3.1 t 検定のロバストネス

次に仮説検定の問題について考えよう。平均 θ 、分散 σ^2 の正規分布に従う

観測値 X_1, \dots, X_n に基づき、仮説 $\theta = \theta_0$ を有意水準 α で検定する場合、 σ^2 が未知のときは t 検定が用いられる。すなわち、不偏分散を s^2 として

$$(7.27) \quad t = \sqrt{n} |\bar{X} - \theta_0| / s$$

を計算し

$$|t| > t_{\alpha/2}(n-1)$$

のとき仮説を棄却する。ただし、 $t_{\alpha/2}(n-1)$ は自由度 $n-1$ の t 分布の上側 $100 \cdot \alpha/2$ パーセント点である。

分布の形 いま、正規分布の仮定が正しくなく、 X_1, \dots, X_n が平均 θ 、分散 σ^2 の正規分布ではない分布に従うとき、仮説 $\theta = \theta_0$ のもとでの t の分布がどうなるかを考えよう。

n がある程度大きければ、中心極限定理により、この場合も \bar{X} はほぼ平均 θ 、分散 σ^2/n の正規分布に従う。また、大数の法則により s^2 はほぼ σ^2 に、また s はほぼ σ に等しくなる。したがって、 t の分布は仮説 $\theta = \theta_0$ のもとでほぼ平均 0、分散 $1/nI(\theta)$ の正規分布になると考えられる。他方、自由度が大きいときの t 分布はほぼ正規分布に一致するから、(7.27) の t の分布は大体 t 分布で近似できると考えてよい。

実は、より詳しい次のような結果が知られている^{*)}。すなわち、眞の分布が正規分布よりすそが長い場合の両側検定については、検定統計量が有意となる点が $\sqrt{3}$ より大きければ(通常用いられる 5%, 1% 検定ではこの条件は満たされる)、仮説 $\theta = \theta_0$ が正しいとき

$$P(|t| > t_{\alpha/2}(n-1)) < \alpha$$

が成り立つ。つまり、第 1 種の過誤の確率は有意水準 α よりむしろ小さくなる。しかも、 n がある程度大きければそれはほぼ α に一致する。

したがって、正規分布の仮定が成り立たなくても、 t 検定を用いれば、誤つて正しい仮説を棄却する確率は α を大きく越えることはないと考えてよい。つまり、 t 検定を用いても結論の妥当性は失われないということになる。このことを、 t 検定は分布の形について妥当性に関してロバストであるという。

*) 卷末柴田(1981)。

検出力 しかし、このことはいつも t 検定を用いてよいということを意味するものではない。 t 検定の検出力は他の検定方式より著しく劣る可能性がある。すなわち、検出力に関して t 検定はロバストでない。そのことを、簡単のため σ^2 を既知として、調べてみよう。

$\sigma^2=1$ と仮定してよい(x_i のかわりに x_i/σ を考えればよい)。正規分布の仮定のもとでの検定方式は、有意水準を α とすれば、(7.3) で述べたように

$$|\bar{X} - \theta_0| \geq z_{\alpha/2} / \sqrt{n}$$

のとき仮説を棄却することになる。

これに対し正規分布の仮定が成り立たず、誤差分布の密度関数が(7.8)であるとしよう。このとき θ の最尤推定量 $\hat{\theta}_{ML}$ は、4.2 節によって n が大きいとき、平均 θ 、分散 $1/nI(\theta)$ の正規分布にほぼ従うものと考えてよいから、 $\hat{\theta}_{ML}$ を標準化すればわかるように

$$(7.28) \quad |\hat{\theta}_{ML} - \theta_0| \geq z_{\alpha/2} / \sqrt{nI(\theta)}$$

のとき仮説を棄却することにすれば、近似的な有意水準 α の検定方式となる。

また、中心極限定理によれば

$$P(|\bar{X} - \theta_0| \geq z_{\alpha/2} / \sqrt{n}) \doteq \alpha$$

なので、 \bar{X} を用いる検定もほぼ正しい(つまり、同じ有意水準)結論を与える。

さて、有意水準をそろえたので、検出力で優劣を比較しよう。対立仮説 $\theta = \theta_1 (\theta_1 \neq \theta_0)$ のもとでは、 $\hat{\theta}_{ML}, \bar{X}$ はほぼ平均 θ_1 、分散がそれぞれ $1/nI(\theta), 1/n$ の正規分布に従うものとみなしてよい。したがって、検定(7.3)の検出力は、帰無仮説からのずれを

$$\delta = \theta_1 - \theta_0$$

として

$$(7.29) \quad \begin{aligned} & P(\sqrt{n} |\bar{X} - \theta_0| \geq z_{\alpha/2}) \\ & = P(\sqrt{n} (\bar{X} - \theta_0) \leq -z_{\alpha/2} \text{ または } \sqrt{n} (\bar{X} - \theta_0) \geq z_{\alpha/2}) \\ & = P(\sqrt{n} (\bar{X} - \theta_1) \leq -z_{\alpha/2} - \sqrt{n} \delta) + P(\sqrt{n} (\bar{X} - \theta_1) \geq z_{\alpha/2} - \sqrt{n} \delta) \\ & = \Phi(-z_{\alpha/2} - \sqrt{n} \delta) + 1 - \Phi(z_{\alpha/2} - \sqrt{n} \delta) \end{aligned}$$

となる。ここで、よく知られているように

$$(7.30) \quad \Phi(z) = \int_{-\infty}^z \phi(x) dx$$

は標準正規分布の累積分布関数である。同様に、検定(7.28)の検出力は

※大学生の常識の範囲とは

筆者の一人は、学期試験で次のような一問題を出したことがある。なお、電卓の持込みを許し、解答時間は他の問題との関係で平均9分である。

〈問〉 日本、フランス、アメリカ合衆国、(旧)ソ連、中国の5国について、国土面積と人口はどのように相関しているか、常識を前提として可能な相関係数を計算せよ。(10点)

出題者の意図は次の通りであった。各国の順位は、

人口： 中国、(旧)ソ連、アメリカ合衆国、日本、フランス

面積： (旧)ソ連、中国、アメリカ、フランス、日本

となるのはほぼ常識であり、それ以上は常識の範囲を超える。多少の異同は問わない。となれば、ノンパラメトリックな相関係数がすぐに計算できる。スピアマン、およびケンドールの順位相関係数^{*} r_s, r_k の両方(あるいは一方)を求めるのに、9分あれば十分である。

面白いのは、「私は常識がないので解答は放棄します」とか、「常識では、ソ連の人口は〇〇億であるから…」という解答であった。〇〇は、その場で自分で考え出した数値である。いずれも、統計学の常識とは、具体的な数値の記憶と考えられているらしい。だが、そのようにして、通常の相関係数(ピアソンの積率相関係数) r を計算しても、日本とソ連の面積は2桁程度異なるので、計算は手間と時間がかかる上に、求めた r の信頼性も低い。

こんな場合に、順位相関のような、値の分布を知らないてもよい(分布によらないdistribution-freeといいう)統計的方法がノンパラメトリック的手法である。ある程度、ゆるい(粗い)常識的なデータに対しても用いることができるという点で、統計学の精神を表しているといえよう。

なお、念のために、人口と面積の数値を挙げておく。

	日本	フランス	アメリカ合衆国	(旧)ソ連	中国
人口	121,492	55,392	241,596	280,144	1,072,218(千人)
国土面積	377,787	547,026	9,363,123	22,402,000	9,596,691(km ²)

(人口は1986年)

*） 本シリーズI『統計学入門』54頁参照。

$$(7.31) \quad P(\sqrt{n I(\theta)} |\bar{X} - \theta_0| \geq z_{\alpha/2}) \\ = \Phi(-z_{\alpha/2} - \sqrt{n I(\theta)} \delta) + 1 - \Phi(z_{\alpha/2} - \sqrt{n I(\theta)} \delta)$$

となる。 $I(\theta)=2>1$ であるから、(7.29)より(7.31)の方が明らかに大きい。実はこの違いは大きい。

例7.4 〈正規性が成立しない場合の検出力の比較〉 誤差分布が(7.8)のとき、仮説 $\theta=0$ を有意水準5%で検定するものとする。標本数は $n=17$ 、対立仮説は $\theta=0.6$ としよう。

\bar{X} による検定の検出力(7.29)は、 $z_{\alpha/2}=1.96$ より

$$(7.32) \quad \Phi(-1.96 - \sqrt{17} \cdot 0.6) + 1 - \Phi(1.96 - \sqrt{17} \cdot 0.6) \\ = \Phi(-4.434) + 1 - \Phi(-0.514) = 0.696,$$

また、 $\hat{\theta}_{ML}$ による検定の検出力(7.31)は

$$\Phi(-1.96 - \sqrt{17 \cdot 2} \cdot 0.6) + 1 - \Phi(1.96 - \sqrt{17 \cdot 2} \cdot 0.6) \\ = \Phi(-5.46) + 1 - \Phi(-1.539) = 0.9381$$

となって、 \bar{X} によるものよりはるかに大きい。

7.3.2 ノンパラメトリック検定

誤差分布の形を特定しなくとも、それが原点について対称であると仮定できるときは、 θ が観測値 X_i の分布の中央値になるので、以下のようにして仮説 $\theta=0$ を検定できる。すなわち

$$(7.33) \quad p = P(X_i \geq 0)$$

とすると、仮説 $\theta=0$ は $p=1/2$ と同等になる。そして

$$N = \{X_1, \dots, X_n\} \text{ のうち, } X_i \geq 0 \text{ のものの個数}$$

とすると、 N は二項分布 $Bi(n, p)$ に従う。したがって、 $p=1/2$ のとき

$$(7.34) \quad P(|N - n/2| \geq c) = \alpha$$

となるように定数 c を定め、 $|N - n/2| \geq c$ のとき仮説を棄却することにすれば、有意水準 α の検定方式が得られる。

たとえば $n=17$ のとき、二項分布 $Bi(17, 1/2)$ において

$$P(N \leq 4) = ({}_{17}C_0 + {}_{17}C_1 + {}_{17}C_2 + {}_{17}C_3 + {}_{17}C_4) / 2^{17} \\ = (1 + 17 + 136 + 680 + 2380) / 131072 = 3214 / 131072 = 0.0245$$

となるから

$$N \leq 4 \text{ または } N \geq 13$$

を棄却域とすれば、有意水準は

$$0.0245 \times 2 = 0.0490$$

となり、ほぼ有意水準 5% の検定方式が得られる。ただし、 N のとり得る値が離散的であるため、(7.34)の左辺が正確に α に等しくなるように棄却域を決めることは一般にはできない。

このような検定方式は、分布が対称であればその形に関係なく、仮説が正しいときに仮説を棄てる確率がつねに α になり、正しい結論を与える。このように分布の形に関係なく妥当な結論を与えるような検定方式をノンパラメトリック検定 non-parametric test という。これに対し、 t 検定のように分布の形について正規分布のような厳密な形を仮定して導かれる検定をパラメトリック検定 parametric test とよぶことがある。上に述べた検定方式はノンパラメトリック検定の一種であるが、観測値の符号のみに注目して結論を導くので一般に符号検定 sign test といわれる。

ノンパラメトリック検定の検出力 ノンパラメトリック検定は、分布の形が正確にわかっているときにはパラメトリック検定よりも検出力が劣る。しかし、分布の形がわからないときには、仮説が正しいときに正確な結論を与えるばかりでなく、仮説が正しくないときの検出力もパラメトリックな検定より大きくなる、すなわち、検出力に関してロバストということも少なくない。

たとえば、誤差分布が(7.8)のとき、 \bar{X} を用いて仮説 $\theta=0$ を有意水準 5% で検定する場合、標本数 $n=17$ 、対立仮説が $\theta=0.6$ のときの検出力は 0.696 であることを(7.32)に示した。(7.32)を導くのに用いたのは、 \bar{X} が仮説のもとで正規分布 $N(0, 1/\sqrt{17})$ 、対立仮説のもとで $N(0.6, 1/\sqrt{17})$ にほぼ従うという事実だけであるから、誤差分布が標準正規分布(7.9)のときも検出力は同じ値 0.696 になる。

一方、符号検定の検出力は真の分布に依存する。真の分布が正規分布(7.9)ならば、 $\theta=0.6$ のときの(7.33)の p の値は

$$p = P(X_i \geq 0) = P(X_i - 0.6 \geq -0.6) = 1 - \Phi(-0.6) = 0.7257$$

となる。 $n=17$ のときの符号検定の有意水準 5% の棄却域は、(7.35)から

$N \leq 4$ または $N \geq 13$ である。したがって、この場合の検出力は、 N が二項分布 $Bi(17, 0.7257)$ に従うときの

$$P(N \leq 4) + P(N \geq 13) = 0.000 + 0.482 = 0.482$$

となる。この値は \bar{X} を用いたときの検出力 0.696 より小さい。

これに対し、真の分布が(7.8)ならば

$$\begin{aligned} p &= P(X_i \geq 0) = P(X_i - 0.6 \geq -0.6) \\ &= \int_{-0.6}^{\infty} \frac{2dx}{\pi(1+x^2)^2} = \frac{1}{\pi} \left[\left(\frac{x}{1+x^2} + \tan^{-1}x \right) \right]_{-0.6}^{\infty} \\ &= \frac{1}{2} + \frac{1}{\pi}(0.4412 + 0.5404) = 0.8125 \end{aligned}$$

となる。したがって、この場合の検出力は、 N が二項分布 $Bi(17, 0.8125)$ に従うことから

$$P(N \leq 4) + P(N \geq 13) = 0.000 + 0.799 = 0.799$$

となる。この値は \bar{X} を用いたときの検出力 0.696 より大きい。

2標本問題のノンパラメトリック検定 分布の形を仮定しないで仮説を検定するノンパラメトリック検定には、状況に応じていろいろな方法がある。ここでは2標本問題の検定について述べよう。第1標本 X_1, \dots, X_m が互いに独立に確率密度関数 $f(x)$ をもつ連続分布に従い、第2標本 Y_1, \dots, Y_n が互いに独立に密度関数 $g(x)$ をもつ連続分布に従うものとする。このとき、2つの分布が等しいという仮説

$$H_0 : f(x) = g(x)$$

を検定することを考えよう。ただし、密度関数 $f(x), g(x)$ の形については何も仮定しないものとする。

いま、2標本 $X_1, \dots, X_m, Y_1, \dots, Y_n$ を一緒にして小さいものから順に並べて、1から $M=m+n$ まで順位 rank をつけ、その中で X_1, \dots, X_m の順位を R_1, \dots, R_m とする。これを X の併合順位 combined rank ということがある。

例 7.5 〈2標本問題の実際〉 2つの標本の比較において、第1標本が $m=10$ で、その値を小さいものから順に並べると、

$$9.5 \ 13.9 \ 18.1 \ 19.9 \ 20.6 \ 21.5 \ 21.8 \ 22.1 \ 25.7 \ 27.9$$

第2標本が $n=14$ で、その値が同じく

14.7 19.8 21.3 21.6 21.7 22.2 22.4 22.7 24.6 27.1
27.0 28.0 29.3 37.1

であった。2つの標本がきたもとの母集団の分布は等しいだろうか。

そこで、これらを一緒にして小さいものから順に並べると

9.5 13.9 14.7 18.1 19.8 19.9 20.6 21.3 21.5 21.6
21.7 21.8 22.1 22.2 22.4 22.7 24.6 25.7 27.1 27.9
27.0 28.0 29.3 37.1

となる(下線は第1標本からのもの)。したがって、 R_1, \dots, R_{10} の実現値は

1, 2, 4, 6, 7, 9, 12, 13, 18, 20

となる。一見して、第1標本、第2標本は互いに他方へ入り込んで混っており、2つのもとの母集団の位置は異なるように見えない。このことは、 R_1, R_2, \dots, R_{10} の値が相対的に大きい方へも小さい方へも偏っていないことからわかる(結論は次例で与えられる)。

さて、仮説 $f(x)=g(x)$ が正しければ、 $X_1, \dots, X_m, Y_1, \dots, Y_n$ はすべて同一の分布に従うから、 R_1, \dots, R_m が $\{1, \dots, M\}$ の中の任意の m 個の数 r_1, \dots, r_m をとる確率はすべて等しく

$$(7.35) \quad P(R_1=r_1, \dots, R_m=r_m) = \frac{1}{M(M-1)\cdots(M-m+1)}$$

となる。そこで、検定統計量として、第1標本の順位和 rank sum

$$(7.36) \quad W = \sum R_i$$

を考えると、仮説のもとで $P(W=w)$ となる確率は分布の形に無関係に計算することができる。

対立仮説のもとで2つの分布が図7.5のように互いにずれているものとすれば、そのズレの方向に応じて、 R_i は小さな値か大きな値をとりやすくなる。

したがって、 W の値も小さな値か大きな値をとりやすくなる。そこで、仮説のもとで

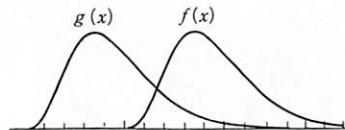


図7.5 2標本問題: ズレのモデル
 $f(x)$ と $g(x)$ はたがいにずれている。ただし、 $f(x), g(x)$ について「形」が同じかどうか前提とされない。重要なことはずれているか(本図)否かということである。

$$(7.37)$$

$$P(W \leq c_1) + P(W \geq c_2) = \alpha$$

となるように c_1, c_2 を定め、 $W \leq c_1$ または $W \geq c_2$ のときに仮説を棄却することにすれば、有意水準 α の検定方式が得られる。これを順位和検定 rank-sum test という。これはまたウィルコクソン検定 Wilcoxon test, あるいはマン-ホイットニー検定 Mann-Whitney test ともよばれる。

定数 c_1, c_2 については、あまり大きくない m, n については表7.2のような数表がつくられている。 m, n がある程度大きければ、仮説が正しいときの W が、平均と分散がそれぞれ

$$(7.38) \quad E(W) = m(M+1)/2, \quad V(W) = mn(M+1)/12$$

の正規分布にほぼ従うことが知られているから

表7.2' 順位和検定の下側 2.5 パーセント点

$n \backslash m$	1	2	3	4	5	6	7	8	9	10
1	—									
2	—	—								
3	—	—	—							
4	—	—	—	10						
5	—	—	6	11	17					
6	—	—	7	12	18	26				
7	—	—	7	13	20	27	36			
8	—	3	8	14	21	29	38	49		
9	—	3	8	14	22	31	40	51	62	
10	—	3	9	15	23	32	42	53	65	78
11	—	3	9	16	24	34	44	55	68	81
12	—	4	10	17	26	35	46	58	71	84
13	—	4	10	18	27	37	48	60	73	88
14	—	4	11	19	28	38	50	62	76	91
15	—	4	11	20	29	40	52	65	79	94
16	—	4	12	21	30	42	54	67	82	97
17	—	5	12	21	32	43	56	70	84	100
18	—	5	13	22	33	45	58	72	87	103
19	—	5	13	23	34	46	60	74	90	107
20	—	5	14	24	35	48	62	77	93	110

2標本問題に対する順位和検定(ウィルコクソン検定、マン-ホイットニー検定)において、順位和の棄却点はこの表から読みとれる。

2組の標本の個数のうち、どちらか大きくない方を m 、他方を n とする。上側 2.5 パーセント点は上で求めた値を $m(m+n+1)$ から引くことで得られる。大きい m, n に対しては正規近似を行う。

$$(7.39) \quad \left| W - \frac{m(M+1)}{2} \right| \geq z_{\alpha/2} \sqrt{\frac{mn(M+1)}{12}}$$

のとき仮説を棄却することにすればよい。

例 7.6 <順位和検定> 直前の例 7.5 では、順位和は

$$W = \sum R_i = 1 + 2 + \cdots + 20 = 92$$

となる。表 7.2 から、 $m=10$, $n=14$ のときの下側 2.5 パーセント点は 91 である。上側 2.5 パーセント点は、表の下の注意から $10(10+14+1)-91=159$ となる。 $91 < 92 < 159$ だから、有意水準 5% で仮説 $H_0: f(x)=g(x)$ は棄却できない。

一方、(7.39)による棄却域の近似限界は、 $E(W)=10(24+1)/2=125$, $V(W)=10 \cdot 14(24+1)/12=291.67$ で、 $z_{0.025}=1.96$ だから

$$|W-125| \geq 1.96 \sqrt{291.67} = 33.47$$

すなわち

$$W \leq 91 \text{ または } W \geq 159$$

となって、表から求めた正確なものと一致する。

順位和検定は、分布の形がどんなものであっても、仮説が正しいときそれを棄却する確率が有意水準に等しくなるという意味で、正しい結論を与える。

7.4 正規分布の仮定のチェック

これまでの節で明らかになったように、データが正規分布からのものであることを仮定して得られた推定や検定の方法は、場合によって非常に悪いものになる。推定ではロバスト推定、検定ではノンパラメトリック検定を用いれば、これらの欠点はかなり補われる。しかしながら、真の分布が正規分布またはそれにきわめて近い分布であるときにそれらの方法を用いれば、多かれ少なかれ効率の低下が起こる。

そこで、最初にデータが正規分布からのものかどうかをチェックして、その結果が肯定的であればパラメトリックな方法を用い、正規分布からのものとはみなせないときにノンパラメトリックな方法を用いたり、あるいは別の分布形を想定するといった方式も考えられる。

このようなチェックには、正規性の検定と正規確率プロットによる図的方法を用いることができる。正規性の検定は、通常の統計的検定と同様に、データが正規分布からのものであるという仮説をたて、仮説が正しくないときに大きい値や小さい値をとりやすいような統計量を用いて、それが棄却域に落ちたときに仮説を棄却するという方式でこのチェックをおこなう。図的方法は正規確率紙にデータをプロットして分布の正規性を視覚的に判断するものである。

7.4.1 正規性の検定

観測値 X_1, \dots, X_n に基づいて、仮説

$H_0: X_1, \dots, X_n$ は正規分布に従うを検定するのが正規性の検定であり、その方法はいくつか提案されている。一般に確率密度関数 $f(x)$ をもつ分布の平均を

$$\mu = \int_{-\infty}^{\infty} xf(x) dx$$

とするとき

$$(7.40) \quad \mu'_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx$$

をこの分布の平均の周りの、 r 次のモーメントとよぶ。いうまでもなく、 $\mu'_2 = \sigma^2$ は分散を表す。

いま、分布 $f(x)$ が平均に関して対称のときは $\mu'_3=0$ だから、これを σ^3 で割って無次元の量にした^{*)}

$$(7.41) \quad \beta_1 = \mu'_3 / \sigma^3$$

を分布の非対称性の指標とすることができます。 β_1 を歪度 skewness とよぶ。正規分布は対称であるから $\beta_1=0$ である。

同様に、 μ'_4 / σ^4 は分布のすその長さとモード周辺の尖り方を表すが、正規分布では $\mu'_4 / \sigma^4 = 3$ となることから、あらたに^{**)}

$$(7.42) \quad \beta_2 = \mu'_4 / \sigma^4 - 3$$

をすその長さの指標として、これを尖度 kurtosis とよぶ。 β_2 は正規分布よりすその長い分布では正の値をとる。

*) , **) 第 I 卷では β_3, β_4 としたが、本卷では正規性の 2 指標という意味で β_1, β_2 と記す。

表7.3 b_1 と b_2 の上側パーセント点

n	α	.05		.01	
		.05	.01	.05	.01
10	1.0	1.4	.9	2.0	
15	.9	1.3	1.1	2.3	
20	.8	1.2	1.1	2.3	
25	.7	1.1	1.1	2.3	
30	.66	.99	1.12	2.21	
40	.59	.87	1.05	2.00	
50	.54	.78	.99	1.87	
75	.45	.66	.87	1.59	
100	.39	.57	.77	1.39	

(出典: 竹内・大橋『統計的推測—2 標本問題』)

正規分布の検定は、標本歪度 b_1 、標本尖度 b_2 によって行うことが多いが、正規性の仮説の棄却の限界となるパーセント点(棄却点)は、この表から読みとれる。

さて、(7.41), (7.42) は母集団の母歪度、母尖度であるが、(7.40) に対応する平均の周りの標本 r 次モーメントを

$$(7.43) \quad m'_r = \sum (X_i - \bar{X})^r / n$$

とすると標本歪度と標本尖度は

$$(7.44) \quad b_1 = m'_3 / (m'_2)^{3/2}$$

$$(7.45) \quad b_2 = m'_4 / (m'_2)^2 - 3$$

と定義される。

観測値の分布が正規分布の場合は b_1, b_2 の期待値はほぼ 0 となるから、それらの絶対値が大きいとき、正規性の仮説を棄却すればよい。それらのパーセント点を表7.3に示す。なお、 n がある程度大きいとき、 b_1, b_2 はほぼ分散 $6/n, 24/n$ の正規分布に従うものとみなすことができる。

例7.7 〈分布の正規性のチェック〉 ポルトの締付けトルクの例7.1では、
 $m'_2=344.30, \quad m'_3=3257.79, \quad m'_4=549848.19$

から

$$b_1=0.510, \quad b_2=1.638$$

が得られる。表7.3から、 $n=30$ のときの b_1, b_2 の5%点はそれぞれ 0.66, 1.12 であり、 b_1 はこの有意水準で有意でないが b_2 は有意である。したがって、締付けトルクの分布は正規分布ではないという結論になる。

また、ラットの肺重量の例7.2では

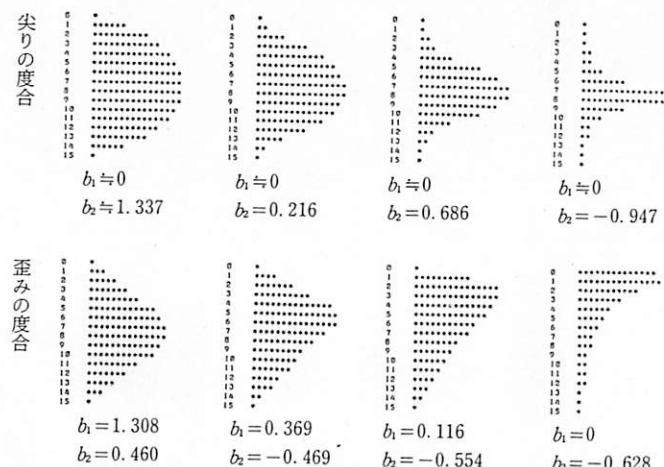
$$m'_2=1.49 \times 10^{-3}, \quad m'_3=8.38 \times 10^{-5}, \quad m'_4=1.206 \times 10^{-5}$$

※パソコンで手軽に見える分布の歪みや尖り※

正規分布のチェックに、分布形の歪み(歪度)、尖り(尖度)を利用することは普通の方法である。したがって、正規分布でなければ、どれほど歪んだり、尖ったり(あるいは尖らなかったり)するものであるか、いろいろと目で見ておくのもよいであろう。ただし、分布を関数として解析的に記述するのは一般的に難しいし、慣れないと手間がかかる。

もし、分布がどれほど歪んでいるか等々を概観的に見たり、計算したければ、思いきってそれを*(アステリック、スター印)で文字入力して画面に出し、他方それをデータのように考えて、標本歪度、標本尖度を計算すればよい。けっこう手軽に分布について学習できる。下図は、BASICによる結果である。

第1の系列は尖度を変えた系列、第2の系列は歪度を変えた系列である。



(出典: 石田『データ解析の基礎』森北出版)

から

$$b_1=1.460, \quad b_2=2.447$$

となる。表7.3は $n=19$ を含まないが、 $n=20$ の棄却点から、これらがともに有意水準 1% で有意であることがわかる。したがって、肺重量の分布も正規分布とはいえない。

7.4.2 正規確率プロットによるチェック

観測値 X_1, \dots, X_n が正規分布 $N(\mu, \sigma^2)$ に従う場合、 $Z_i = (X_i - \mu)/\sigma$ の分布は標準正規分布になるので、 X_i が特定の値 x 以下になる確率 p_x は、(7.30) の標準正規分布の累積分布関数 $\Phi(z)$ を用いて

$$p_x = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

と表され、これは Φ の逆関数を Φ^{-1} とすると

$$(7.46) \quad \Phi^{-1}(p_x) = \frac{x - \mu}{\sigma}$$

と同等である、(7.46) は、 $\Phi^{-1}(p_x)$ が x の 1 次式になることを示す。

そこで、観測値を大きさの順にならべた順序統計量を

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

とし、グラフ用紙に

$$(7.47) \quad (X_{(i)}, z_i), \quad z_i = \Phi^{-1}(i/(n+1)), \quad i=1, 2, \dots, n$$

をプロットすれば、正規分布の仮定が正しいときの打点はほぼ直線的になることが期待される。これが正規確率プロットであり、そのプロットを容易に行うためのグラフ用紙が正規確率紙 normal probability paper (次頁参照) として市

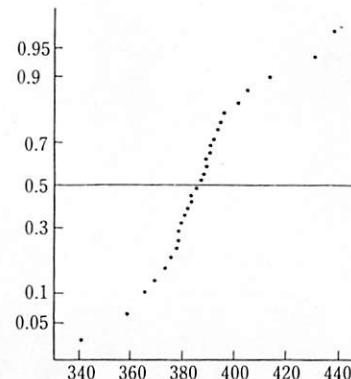


図 7.6(a) 正規確率プロット
(ボルトの締付けトルクのデータ)

打点が両端で傾きが小さくなる S 字形の点線となっている。正規分布よりそその長い分布からのデータに典型的である。

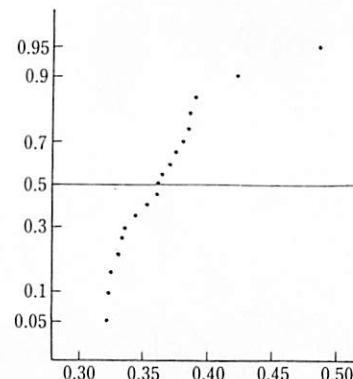
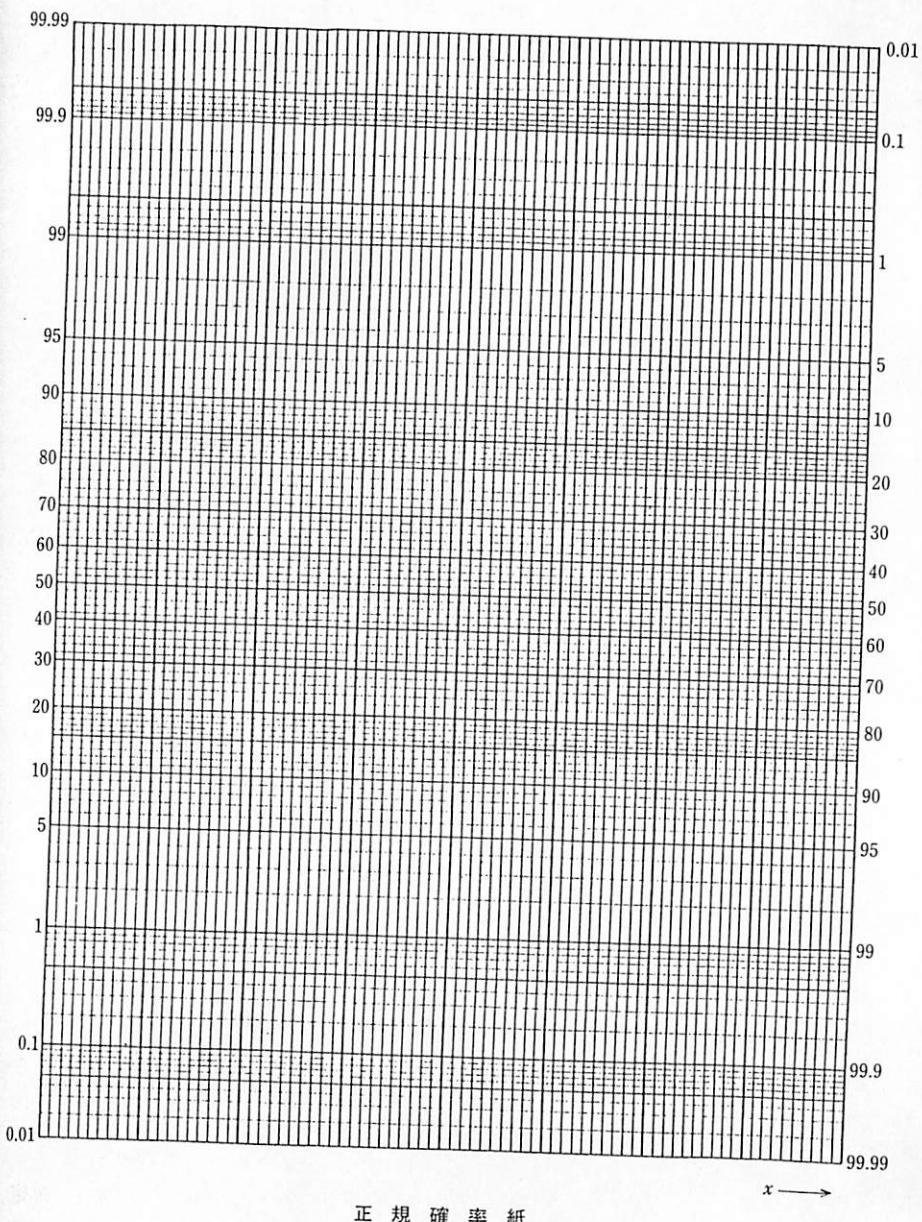


図 7.6(b) 正規確率プロット
(ラットの肺重量のデータ)

右にそそを引く非対称な分布からのデータに典型的に見られる形をしている。



正規確率紙

販されている。なお、プロット位置 z_i については(7.47)の他にもいくつか提案されている。

図7.6(a)はボルトの締付けトルクの例7.1の、同じく図7.6(b)はラットの肺重量の例7.2の正規確率プロットを示す。図7.6(a)では、打点は両端で傾きが小さくなるS字形の曲線状になっている。これは正規分布よりその長い分布からのデータに典型的である。図7.6(b)では、打点は上に凸な曲線にみえる。これは右にすそをひく非対称な分布からのデータでよくみられるものである。

◆◆◆◆ 第7章 練習問題 ◆◆◆◆

7.1 <コーシー分布の最尤推定量と標本中央値の比較> 最尤推定量と標本中央値を比較するために

i) コーシー分布

$$f_{\theta}(x) = \frac{1}{\pi(1+(x-\theta)^2)}, \quad -\infty < x < \infty$$

のフィッシャー情報量 $I(\theta)$ を求めよ。標本の大きさ n が十分大きいときの θ の最尤推定量 $\hat{\theta}_{ML}$ の分散 $V(\hat{\theta}_{ML})$ はいくらくか。

ii) 標本の大きさ n が十分大きいときの標本中央値 X_{med} の分散 $V(X_{med})$ はいくらくか。

iii) $V(X_{med})$ の $V(\hat{\theta}_{ML})$ に対する比を求めよ。

7.2 <標本平均と標本中央値>

i) 観測値 X_1, X_2, \dots, X_n が与えられたとき、

$$\sum_{i=1}^n (X_i - \theta)^2$$

を最小にする θ は $\theta = \bar{X}$ であることを示せ。

ii) 同じく $\theta = X_{med}$ は

$$\sum_{i=1}^n |X_i - \theta|$$

を最小にすることを示せ。

(注) ii) は順序統計量 $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ を考えれば、扱い易くなる。

7.3 <最尤推定量> 下のデータに分布(7.8)をあてはめたときの θ の最尤推定量を、本文に述べた方法で求めよ。中央値を最初の近似解 $\theta^{(0)}$ として、逐次近似解

$\theta^{(k)}$ と係数 $W_1^{(k)}$ の収束の様子を観察せよ。

8.2 7.5 8.7 8.4 9.6

7.4 <幹葉表示とロバスト推定> 下のデータはプラスチック製品の成型工程で製品の肉厚(mm)を測定したものである。

4.95	5.02	5.08	4.90	5.12	5.19	4.80	4.87	4.98
4.58	4.81	5.18	5.25	5.05	4.79	5.57	5.01	5.57
5.01	4.81	5.13	5.04	5.13	5.16	4.69		

- i) このデータを幹葉表示せよ。
- ii) 標本平均 \bar{X} を求めよ。
- iii) 標本中央値 X_{med} を求めよ。
- iv) $k=3$ のトリム平均 $\hat{\theta}_{trim}(3)$ を求めよ。
- v) フーバー型推定量 $\hat{\theta}_H$ の値を求めよ。
- vi) 以上の推定値を比較して考察せよ。

7.5 <対称性の符号検定> 次のデータが対称な分布から得られたものとして、符号検定を用いて、平均が0であるという仮説を有意水準5%で両側検定せよ。

5.8	0.1	1.9	-2.5	5.7	-1.2	4.2	8.0	7.9	-2.3
0.9	5.5	3.3	-7.2	1.2	-5.4	3.1			

7.6 <2標本問題と順位和検定> 2つの標本がある。第1標本の標本の大きさは15で、その値を小さいものから順に並べると、

42.6	43.9	44.1	44.2	45.0	45.8	47.4	48.0	48.7
49.5	49.6	50.0	51.4	52.7	52.8			

であり、第2標本の標本の大きさは10で、その値は、同じく

44.7	46.8	46.9	47.1	48.2	50.0	52.1	53.5	54.2	56.3
------	------	------	------	------	------	------	------	------	------

である。順位和検定により、2つの分布が等しいという仮説を有意水準5%で検定せよ。

7.7 <正規性の仮定のチェック> 次のデータ(A), (B)について、それぞれ

- i) 正規確率紙へのプロットを行い、考察せよ。
- ii) 標本歪度 b_1 および標本尖度 b_2 を求めよ。データは正規分布からのものとみなせるか。

(A)	10.5	10.4	10.6	11.9	11.3	11.8	11.7	12.6	10.9
	11.1	13.0	11.2	13.4	10.8	11.2			
(B)	20.0	18.3	20.3	37.1	20.2	21.0	6.5	24.0	14.2
	21.0	18.3	13.4	14.8	21.6	16.8			

第8章 質的データの統計的分析

統計学といえば、ふつうは‘量の統計学’と考えられている。身長、血圧、物価、収入、得点、支持率などみな「量」である。それと対照的に、買う買わない、合格しない、有効無効、投票するしない、(医学・生物学での)生きる死ぬなどは、「質」的変数である。プロビット、ロジット・モデルは、これらを扱う有用なアプローチとして発展してきたものであり、ほぼ確立した方法として定着している。本章では、理論から応用にわたって最重要ポイントを説明する。

8.1 二値データ

8.1.1 質的データ

分析対象となる事象には数値データを得ることができず、標本の質的特性のみがデータとして得られる場合がある。たとえば、薬剤による生物の生死、強度試験による製品の破壊の有無、消費者が自動車を購入したかどうか、学生がどの科目を選択したか、などでは数値データを得ることはできず、得られるデータは生物の生死、破壊の有無、購入の有無などの二項反応 binary response (一般には、多項選択ないしは多項反応)である。このように、定量化できず、対象がある状態やカテゴリーの項目に属しているかどうかのみを知ることができるデータを質的データとよぶ。表8.1はその例である。

今まで主に数値をとる変数を考えてきたが、質的データを扱えないのでは、統計分析の有用性はかなり限られたものとなってしまう。この分析のためには、数多くの統計的な手法が考えられており、実際の分析に広く利用されている。その1つは分割表に基づく方法であるが、いま1つは、回帰分析の考え方を応用した確率モデルを使う方法である。

表8.1 薬物濃度と生物の死亡数(付表参照)

	$\log_2(\text{濃度})$				
	0	1	2	3	4
死 亡 数	2	8	15	23	27
個 体 数	30	30	30	30	30
死 亡 数 (%)	6.7	26.7	50.0	76.7	90.0

(出典:D.R.コックス著、後藤昌司他(訳)『二値データの解析』)

付表 個別ケース別の対数濃度 X_i 、死亡・生存の別

Y_i (表8.1参照)

個体	31	32	...	37	38	39	40
X_i	1	1	...	1	1	1	1
Y_i	1	1	...	1	1	0	0
個体	•	59	60	61	62	•	•	75	76	77
X_i	•	1	1	2	2	•	•	2	2	2
Y_i	•	0	0	1	1	•	•	1	0	0
個体	•	•	89	90	91	92	•	•	•	•
X_i	•	•	2	2	3	3	•	•	•	•
Y_i	•	•	0	0	1	1	•	•	•	•
個体	116	117	118	119	120					
X_i	3	3	3	3	3					
Y_i	1	1	0	0	0					

質的データの中で最も基本的であり、かつ重要であるのは、観測結果が成功、失敗のように2つの状態をとる二項的 binary な場合である。この場合、一方の状態をとるとき(成功の場合)1、他方の状態をとるとき(失敗の場合)0のダミー変数 dummy variable Y_i を使って観測結果を表すが、このようなデータは二値データ binary data とよばれている。二値データの統計分析手法で最も広く使われているプロビット・モデル probit model、ロジット・モデル logit model(ロジスティック・モデル logistic model ともよばれる)を中心に二値データの分析方法について、わかりやすく説明しよう。

量と反応の関係を示している、いわゆる量-反応関係 dose-response relationship といわれるものの典型である。 $X=2$ の周辺が半数が死する半数致死量 lethal dose 50、いわゆる LD50 である。

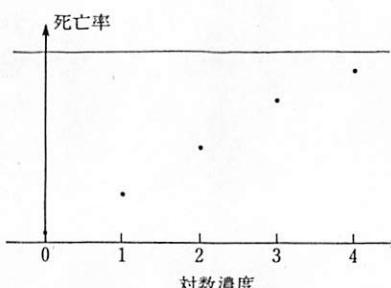


図8.1 薬物の対数濃度と生物の死亡率
半分ずつの稀釈を繰り返すから対数濃度は2を底としている。稀釈とともに死亡率は低下していく。

8.1.2 二値データの確率モデル

結果が表8.1、同付表のように2つの状態をとりうる場合、観測や実験の結果得られるデータが二値データである。観測結果は1つの状態をとる場合0、他の状態の場合1のダミー変数 Y_i で表すことができるが、当然のことながら $Y_i=1$ となる確率を統計的に分析することが必要となる。

$i=1, 2, \dots, 150$ は本来の個体番号で、30~120だけを示す。対数濃度が $X=1, 2, 3$ となるにつれ、 $Y_i=1$ (死亡)となる割合は高進することがわかる。

実際に、ロジット、プロビット・モデルの分析を行うには、この表の形による。なお、表8.1の原典には、この表までは付されていないが、基本的にはこの付表の形になる。

例8.1 <量・反応関係> 薬物の投与水準に基づく生物の死亡数として表8.1のようなデータが得られたとしよう。付表のように、 Y_i を i 番目の個体に対する薬物の効果を表す変数とし、薬物によって生物が死する場合1、生存する場合0としよう。 $Y_i=1$ となる確率(死亡率)は、一定量を単位とした薬物の対数濃度 X_i に影響され、濃度が高くなるに従ってその値は大きくなる。したがって、 $Y_i=1$ となる確率とそれに影響を与えると考えられる変数 X_i との関係を求めるモデルを考え、それに基づいて質的データを分析することが考えられる。この結論は例8.2で与えられる。

そこで、ある X_i の値に対する $Y_i=1$ となる確率 P_i は、 Y_i の X_i に関する条件付確率 $P(Y_i=1|X_i)$ であるが、これは X_i の関数^{*)}であって、

$$(8.1) \quad P(Y_i=1|X_i)=F^*(X_i)$$

と表すことができる。ここで問題となるのは F^* としてどのような関数を使うかである。

最も簡単なものとしては F^* を線形関数

$$F^*(X_i)=\alpha_0+\alpha_1 X_i$$

とすることが考えられる。このようなモデルは線形確率モデル linear probability model とよばれ、推定が簡単なため、コンピューターの計算能力が十分

*) F に * を付したのは、後の混同を避けるためである。

発達していない時代にはしばしば使われたが、このままでは X_i の小さな値や大きな値に対しては $F^*(X_i)$ が負になったり、1 を越えてしまうという問題がある。また、これを改良して、負の場合は 0、1 を越える場合は 1 とする方法などが考えられたが、確率が 0 や 1 に近い部分ではモデルの当てはまりが良くないため、現在では線形確率モデルやそれを改良したモデルはほとんど使われていない。

二值データの分析で一般に使われるモデルは回帰分析を応用した次のような考え方に基づくものである。いまここで、 Y_i が 0 をとるか 1 をとるかを決定する仮想的な因子 Y_i^* があり、それが

$$(8.2) \quad Y_i^* = \beta_0 + \beta_1 X_i + \varepsilon_i,$$

で表すことができるとする。 Y_i^* は直接に値を観測できない連続的な変数であるが、 X_i によって体系的に決まる部分 $\beta_0 + \beta_1 X_i$ とそれ以外の個体差に基づく偶然的な部分 ε_i の和で表すことができる。回帰モデルと同じく、 ε_i は誤差項 error term とよばれ、既知の確率分布に従う確率変数とする。

われわれは Y_i^* を直接観測できないが、 Y_i は Y_i^* の符号によって

$$(8.3) \quad Y_i = \begin{cases} 1 & Y_i^* > 0 \text{ の場合} \\ 0 & Y_i^* \leq 0 \text{ の場合} \end{cases}$$

と決定されるものとする。

このモデルはどのように得られるか、表 8.1 の薬物の効果を例にとり考察してみよう。 i 番目の個体に X_i の濃度の薬物が投与されたとしよう。個体 i が死ぬ ($Y_i=1$) か、死なない ($Y_i=0$) かどうかは、対数濃度 X_i に依存する薬剤の効力 E_i という体系的部分と、個々の個体によって異なる耐久力(一般には臨界値 cutoff value) C_i という偶然的部分に依存すると考えられる。 $Y_i=1$ となるのは $E_i > C_i$ のときであり、 $Y_i=0$ となるのは $E_i \leq C_i$ のときである。効力 E_i は薬物濃度 X_i の線形関数として決まり、他方 C_i はある確率分布に従う確率変数とすると、 Y_i は

$$Y_i^* = E_i - C_i$$

の符号から、(8.3) によって決まる。平均的耐久力 $E(C_i)$ の方だけ調整して $Y_i^* = E_i - E(C_i) - (C_i - E(C_i))$ とすると、もともと E_i が X_i の線形関数であるから、これから (8.2) の形

$$Y_i^* = \beta_0 + \beta_1 X_i + \varepsilon_i$$

が得られる。ここで、 $\varepsilon_i = -C_i + E(C_i)$ である。

※「中毒センター」よりの報告※

身の回りにあるいろいろな化合物や薬も、使い方を誤ると思わぬ事態を招くことがある。たとえば、「座薬」を正座して飲んだとか、片仮名を習った来日したばかりの留学生が洗剤の「ママ・レモン」をジュースと誤って飲んだとか、中性洗剤をサラダ・オイルと間違えて御飯をいためて食べたとか、いずれも、現実に報告のあったケースである。注意書きをよく読まなかつたり、使用中に品名ラベルがはがれれば、あってもおかしくない事態である。また、大人がコーラの飲みさしの缶を灰皿替りに用いて、子供が誤って飲むという事故例もある。タバコ中のニコチンは 1 本当たり 0.8 mg であり、ニコチンは 200 mg も摂取すれば致死とされるので、ただちにというわけではないが、体の小さい幼児にとって危険であることに変りはない。

いま、 F を $-\varepsilon_i$ の累積分布関数としよう(多くの場合そうであるように誤差項 ε_i の確率分布が原点 0 に関して対称なら、 F は ε_i 自体の累積分布関数である)。 $Y_i=1$ となるのは、 $-\varepsilon_i \leq \beta_0 + \beta_1 X_i$ のときであるから、(8.1) の、 $Y_i=1$ となる確率は、この F を用いて

$$(8.4) \quad P(Y_i=1|X_i) = F^*(X_i) = F(\beta_0 + \beta_1 X_i)$$

で与えられる。プロビット・モデルでは、(8.1) の F^* として標準正規分布、ロジット・モデルでは、ロジスティック分布 logistic distribution が、(8.4) の

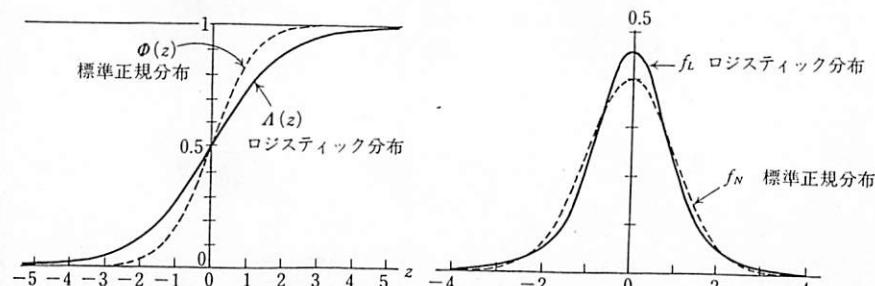


図 8.2 標準正規分布、ロジスティック分布の累積分布関数の比較。

両者とも 0 から 1 まで滑らかに変化し、また、確率分布のメディアン(確率 = 0.5)で、傾きが最大であり、結果の確率に最も大きい影響を与える。他方、両者には多少の差もある。

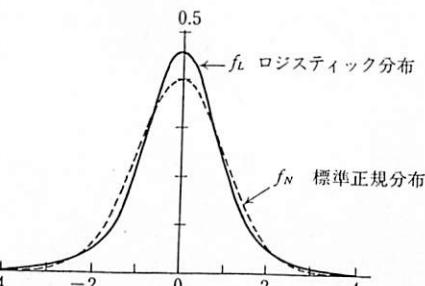


図 8.3 標準正規分布と分散 1 のロジスティック分布の密度関数の比較。

ともに分散を 1 に揃えると、ロジスティック分布 f_L の方が多少分布の広がりが大きいが、その差は小さい。両分布ともに、0 に関して左右対称である。

形で採用される。いずれも、原点に関し対称な確率分布である(図8.2, 8.3)。

8.2 ロジット・モデルとプロビット・モデル

8.2.1 プロビット・モデル

各種の推定や検定においては評差項が正規分布に従うと仮定する場合が多いが、プロビット・モデルは、(8.2)の ε_i が標準正規分布に従うとし、その累積分布関数

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

を使い、(8.1)における $Y_i=1$ となる確率を

$$F^*(X_i) = \Phi(\beta_0 + \beta_1 X_i)$$

とするものである。なお、(8.2)の右辺を正の定数倍してもその符号は変わらないから、一般性を失うことなく $V(\varepsilon_i)=1$ とすることができる。

8.2.2 ロジット・モデル

ロジット・モデルは、誤差項の確率分布がロジスティック分布であるとし、ロジスティック分布の累積分布関数

$$\Lambda(z) = \frac{e^z}{1+e^z}$$

を用いて、 $Y_i=1$ となる確率を

$$F^*(X_i) = \Lambda(\beta_0 + \beta_1 X_i)$$

とするものである。ところで

$$P_{0i} = P(Y_i=0|X_i), \quad P_{1i} = P(Y_i=1|X_i)$$

とすると

$$P_{0i} = 1 - P_{1i} = 1/[1 + \exp(\beta_0 + \beta_1 X_i)]$$

であるから

$$(8.5) \quad \log(P_{1i}/P_{0i}) = \beta_0 + \beta_1 X_i$$

となる。すなわち、ロジット・モデルは Y_i が0となるか1となるかの2つの確率の比(オッズ odds)の対数、つまり対数オッズが X_i の線形関数であると

したものである。

プロビット・モデルとロジット・モデルの比較 二値データが得られた場合、プロビット・モデルとロジット・モデルのどちらを使うかが問題となる。回帰分析の考え方からは(8.2)の誤差項が正規分布であると仮定するプロビット・モデルのほうがより自然であるが、ロジット・モデルにはi)分布を簡単な関数で表すことができるため数学的な取り扱いが容易である、ii)ロジット・モデルは対数オッズが X_i の線形関数となるという利点がある。

次に、正規分布とロジスティック分布を比較してみる。標準正規分布の分散は1であり、ロジスティック分布の分散は詳しい証明は省くが $\pi^2/3$ である。この分散の違いを取り除くために、ロジスティック分布の分散が1となるように正規化するとロジスティック分布の密度関数 $f_L(z)$ は、

$$f_L(z) = \exp(z/d)/[1 + \exp(z/d)]^2 \cdot d, \quad d = \sqrt{3}/\pi$$

となる。これと標準正規分布の密度関数

$$f_N(z) = \exp(-z^2/2)/\sqrt{2\pi}$$

を比較すると図8.3のようになる。

ロジスティック分布の方が多少分布の広がりが大きいがその差は小さい。このため、観測値が $Y_i=1$ となる確率が大きく1に近い場合やその逆に小さく0に近い場合以外は、いずれでもあまり差はない。

ロジット・モデルによる各々の母数の推定値は、プロビット・モデルによる同じ母数の推定値のほぼ $\pi/\sqrt{3} \approx 1.81$ 倍程度の大きさになることがわかっている。

したがって、二値データの分析においては両者の間に決定的な優位関係はあるとは言えないが、モデルを作る場合の考え方(回帰分析的な考え方を主にするか、対数オッズが線形であることを主とするか)や他の分析手法との関連からどちらを使うかが決定される。経済データではプロビット・モデルが、自然科学の分野ではロジット・モデルが使われることが比較的多い。

8.2.3 モデルの推定方法

(8.5)は未知の母数 β_0, β_1 を含んでおり、これを実際のデータから推定しなければならない。この推定には最尤法 maximum likelihood method が今日では最もよく使われている。

$-\varepsilon_i$ の累積分布関数を既知の関数 $F(z)$ としよう。プロビット・モデルの場合、 $F(z)$ は標準正規分布の累積分布関数 $\Phi(z)$ であり、ロジット・モデルの場合、 $F(z)$ はロジスティック分布の累積分布関数 $\Lambda(z)$ である。 $P(Y_i=1|X_i) = F(\beta_0 + \beta_1 X_i)$ および $P(Y_i=0|X_i) = 1 - F(\beta_0 + \beta_1 X_i)$ であるから、ある標

本 Y_1, Y_2, \dots, Y_n を得る同時確率は、 $Y_i=1$ の場合 $F(\beta_0 + \beta_1 X_i)$ を、 $Y_i=0$ の場合 $1 - F(\beta_0 + \beta_1 X_i)$ をあわせて n 個乗じたものである。この確率を β_0, β_1 の関数と考えれば、次のような尤度関数を得ることができる。

$$(8.6) \quad L(\beta_0, \beta_1) = \prod_{Y_i=1} F(\beta_0 + \beta_1 X_i) \cdot \prod_{Y_i=0} [1 - F(\beta_0 + \beta_1 X_i)]$$

ここで記号 \prod は、それぞれ、 $Y_i=1$ となる i についての積、および $Y_i=0$ となる i についての積を表す。

最尤推定量 $\hat{\beta}_0, \hat{\beta}_1$ は $L(\beta_0, \beta_1)$ を最大にするものであるが、この推定量は F を標準正規分布とした場合プロビット最尤推定量 probit maximum likelihood estimator、ロジスティック分布とした場合ロジット最尤推定量 logit maximum likelihood estimator とよばれる。

(8.6) は積の形であるので数学的な取り扱いを容易にするために両辺の対数をとれば

$$(8.7) \quad \log L(\beta_0, \beta_1) = \sum_{i=1}^n \{ Y_i \log F(\beta_0 + \beta_1 X_i) + (1 - Y_i) \log [1 - F(\beta_0 + \beta_1 X_i)] \}$$

となる。 $\log L(\beta_0, \beta_1)$ を最大にする推定量は

$$(8.8) \quad \partial \log L(\beta_0, \beta_1) / \partial \beta_0 = 0, \quad \partial \log L(\beta_0, \beta_1) / \partial \beta_1 = 0$$

の解として求められるが、(8.8) はプロビット、ロジット・モデルのいずれにおいても複雑な関数形であるため、一般的回帰分析の場合と異なり解析的にこの連立方程式の解を求ることはできない。このため実際の推定では、関数の零点を求めるニュートン法などの標準的なアルゴリズムによって、数値計算で β_0, β_1 の値を求めることになるが、プロビット、ロジット・モデルいずれの場合でも、 $\log L(\beta_0, \beta_1)$ は β_0, β_1 の凹関数であるため(8.8) が複数の解をもつことはなく、比較的容易に計算することができる。

プロビット、ロジット・モデルのいずれにおいても(8.8) が複数の解をもつことはない。しかしながら、(8.8) が解をもたず、最尤推定量が存在しない場合がある。図 8.4 のように Y_i が X_i の値がある定数 c より大きい場合すべて 1、小さい場合すべて 0 としよう。すると、 $\log L(\beta_0, \beta_1)$ は図 8.5 のようになり、これが最大になるのは、 β_1 の値が無限大のときである。したがって、このような場合、実用上意味のある推

❖「アッ」という間に解に収束：ニュートン法❖

今日ではニュートンの「方程式の逐次近似解法」(ニュートン法)として知られているものは、どんな大学初級の基礎的な解析学の教科書にものっているものである。

このニュートン法とは、次の定理で保証される方法である。

区間 $[a, b]$ において $f(x)$ が 2 回微分可能で、 $f''(x) > 0$ 、また $f(a)$ と $f(b)$ とは符号が異なるとする。このとき区間 (a, b) 間にただ 1 つ $f(x) = 0$ の根(解) ξ がある。

そこで、 $f(x_1) > 0$ である (a, b) 内の x_1 から始めて、順次 $y = f(x)$ のグラフの $x = x_n$ における接線

$$l : y - f(x_n) = f'(x_n)(x - x_n)$$

を作り、横軸との交点を

$$x_{n+1} = x_n - f(x_n) / f'(x_n)$$

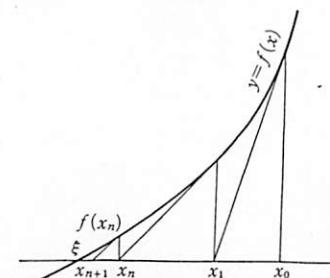
として漸化的に点列 $\{x_n\}$ を定義すると、 $\{x_n\}$ は単調に根 ξ に収束する。

この解法はわかり易い上に、収束がきわめて速く誤差が 2 乗の程度で減少してゆく 2 次の収束性

$$|x_{n+1} - \xi| \leq C|x_n - \xi|^2 \quad (C \text{ は定数})$$

があるといわれている。実際、 $f(x) \equiv x^2 - a (a > 0)$ として、平方根 \sqrt{a} の計算に応用してみると、 $0.01 \leq a \leq 1$ に対して、6 回以下の反復で 10^{-10} 以下の誤差になる。ここで、 $x_1 \equiv 1$ としている。

とにかく速いのである。



ニュートン法(方程式の逐次近似解法)
接線を用いて、根 ξ に収束する点列 $\{x_n\}$ を作る。単調に収束するだけでなく、その収束も非常に速い。



ニュートン(Sir Isaac Newton, 1642-1727)
終生独身であり、神を信じる篤信の人でもあった。

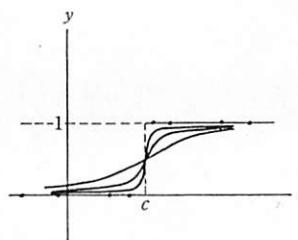


図 8.4

図 8.4 のようにある $x=c$ を境に一方がすべて $Y_i=0$, 他方が $Y_i=1$ となると, $\log L$ は図 8.5 のようになり最尤推定値を計算することができない。すなわち, データに無限によくあてはまる β_1 の値が存在する。

定量を求めることができず, 分析に用いることはできない。

なお, このような現象は標本の大きさ n が大きいときはほとんど起こらないが, n が小さいときは無視できない割合で発生する。実際の推定で, 推定値が極端に大きかったり小さかったり, 繰り返し計算が収束しない場合などは, この現象が起つていなかどうかを調べてみるとよい。

プロビット, ロジット最尤推定量 $\hat{\beta}_0, \hat{\beta}_1$ は一致推定量であり, 標本の大きさ n が増加するに従って β_0, β_1 の真の母数の値に近づく。ただし, 不偏性は一般には成り立たない。また, その標本分布を正確に求めることはむずかしいが, 中心極限定理から n が無限大になった漸近分布は平均 β_0 あるいは β_1 の正規分布である。多くのコンピューター・プログラムでは $\hat{\beta}_0, \hat{\beta}_1$ の漸近分散・共分散は推定値と共に計算され出力される。 $\hat{\beta}_0, \hat{\beta}_1$ の漸近分散・共分散をわかりやすく表すのには行列が必要なためこれについては, 8.4 節において触れる。これによって, 回帰分析と同様の母数に関する, t -分布を使った β_0, β_1 についての仮説検定(自由度は $n-2$)を行うことができる。

プロビット, ロジット・モデルは, 尤度関数(8.7)が複雑であり, β_0, β_1 の推定にはコンピューターの使用が必要不可欠である。今日では, TSP や SAS など多くの標準的な統計パッケージに, これらの推定プログラムが組み込まれており, 簡単にモデルの推定を行うことができる。

例 8.2 〈モデルの推定例：量・反応関係〉 TSP によって表 8.1 の量・

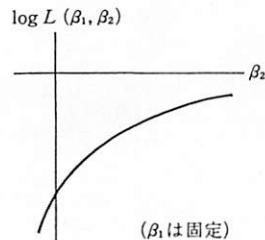


図 8.5

反応関係のデータを使い, プロビット, ロジット・モデルによってモデルの推定を行った。結果は次のようにあった。プロビット・モデルでは

$$Y^* = -1.086 + 0.543X + \varepsilon \quad (8.71) \quad (0.09)$$

となる。()内は標準誤差である。またロジット・モデル^{*)}では

$$Y^* = -2.324 + 1.162X + \varepsilon \quad (0.42) \quad (0.18)$$

となる。(出力例を次頁に示しておいた)

さらに薬物の効果について検定してみよう。帰無仮説は薬物の効果がない。すなわち, $H_0: \beta_1=0$ であり, 対立仮説は $H_1: \beta_1>0$ である。 t -値はプロビット・モデルの場合が 6.090, ロジット・モデルの場合が 6.405 であるから, 常識的な有意水準では帰無仮説は棄却され, この薬物は効果がある。

なお, LOG OF LIKELIHOOD FUNCTION は最大化された対数尤度関数の値であり, これは 8.4 節で述べる尤度比検定に用いられる。

8.3 確率の推定

ふつうは β_0, β_1 の推定がもっとも重要だが, 二値データでは, $Y_i=1$ となる確率 $P_i=P(Y_i=1|X_i)$ の推定値がほしいことも多い。 $P(Y_i=1|X_i)$ は, 最尤推定量 $\hat{\beta}_0, \hat{\beta}_1$ を使い, (8.4) から

$$(8.9) \quad \hat{P}_i = F(\hat{\beta}_0 + \hat{\beta}_1 X_i)$$

で推定する。 \hat{P}_i は確率変数であり, 区間推定や仮説検定を行うことができるが, そのためにはその標本分布を知る必要がある。これは一般的には難しいが, その漸近分布は比較的容易に求めることができる。

\hat{P}_i の漸近分布は正規分布であり, $\hat{\beta}_0, \hat{\beta}_1$ の漸近分散を σ_0^2, σ_1^2 , 漸近共分散を σ_{01} , $f(z)$ を ε_i の密度関数とすると, \hat{P}_i の標本分布は n が十分大きいときは

$$(8.10) \quad N(P_i, f(\hat{\beta}_0 + \hat{\beta}_1 X_i)^2 \cdot (\sigma_0^2 + \sigma_1^2 X_i^2 + 2\sigma_{01} X_i))$$

で近似できる。これを用いて区間推定や検定を行うことができる。

証明は次の通り。いま, $f=F'$ とすればティラー展開によって

*) TSP では, 多項モデルの推定(multinomial logit estimation)も可能である。

量・反応関係のデータに対するロジット・モデル、プロビット・モデルのTSPによる推定(例8.2)

TSP コマンド

```
1 ? smpl 1 150;          1)
2 ? load(file = "A:input.dat") x y;    2)
3 ? probit y c x;        3)
4 ? logit y c x;        4)
```

1) 標本の大きさ n の指定 2) ファイルから入力 3) プロビットを指定 4) ロジットを指定

プロビット・モデルの推定結果

PROBIT ESTIMATION

DEPENDENT VARIABLE: Y

LOG OF LIKELIHOOD FUNCTION =	-72.9066
NUMBER OF OBSERVATIONS =	150
NUMBER OF POSITIVE OBSERVATIONS =	75 0)
PERCENT POSITIVE OBSERVATIONS =	0.500000
SUM OF SQUARED RESIDUALS =	23.6369 1)
R-SQUARED =	0.377395 2)
PERCENT CORRECT PREDICTIONS =	0.766667 3)

STANDARD

PARAMETER	ESTIMATE	ERROR	T-STATISTIC
C	-1.0862	8.7057	-0.12477
X	0.54310	0.89178E-01	6.0901

ロジット・モデルの推定結果

MULTINOMIAL LOGIT ESTIMATION

DEPENDENT VARIABLE: Y

LOG OF LIKELIHOOD FUNCTION =	-71.7798
NUMBER OF CASES =	150
NUMBER OF CHOICES =	300
SUM OF SQUARED RESIDUALS =	23.3434 1)
R-SQUARED =	0.377511 2)
PERCENT CORRECT PREDICTIONS =	0.766667 3)

STANDARD

PARAMETER	ESTIMATE	ERROR	T-STATISTIC
C1	-2.3238	0.41789	-5.5608
X1	1.1619	0.18142	6.4046

0) $Y_i=1$ となった観測値の個数 1) $\sum [Y_i - F(\hat{\beta}_0 + \hat{\beta}_1 X_i)]^2$ 2) $1 - \sum [Y_i - F(\hat{\beta}_0 + \hat{\beta}_1 X_i)]^2 / \sum (Y_i - \bar{Y})^2$ 3) Y_i の予測値を, $\hat{\beta}_0 + \hat{\beta}_1 X_i > 0$ のとき 1, < 0 のとき 0 とする(8.3節参照)

8.3 確率の推定

$$(8.11) \quad \hat{P}_i - P_i = F(\hat{\beta}_0 + \hat{\beta}_1 X_i) - F(\beta_0 + \beta_1 X_i) \\ = f(\beta_0^* + \beta_1^* X_i) \{ (\hat{\beta}_0 - \beta_0) + (\hat{\beta}_1 - \beta_1) X_i \}$$

となる。ただし、 β_0^* , β_1^* はそれぞれ β_0 と $\hat{\beta}_0$, β_1 と $\hat{\beta}_1$ の間の値である。ここで $(\hat{\beta}_0, \hat{\beta}_1)$ が一致推定量であり、かつその漸近分布が正規分布であることを用いる。

例8.3 〈確率の推定の例：量・反応関係〉 表8.1の量・反応関係のデータでは、プロビット・モデルで、

$$\hat{P}_i = \Phi(-1.086 + 0.543 X_i)$$

ロジット・モデルで、

$$\hat{P}_i = A(-2.324 + 1.162 X_i)$$

となる。とくにこの場合は対数オッズが(8.5)に従い、

$$\log \frac{\hat{P}_i}{1-\hat{P}_i} = -2.324 + 1.162 X_i$$

と書けることに注意する。また、 $\hat{P}_i > 1/2$ となるのは、それぞれ

$$-1.086 + 0.543 X_i > 0 \quad (\text{プロビット・モデル})$$

$$-2.324 + 1.162 X_i > 0 \quad (\text{ロジット・モデル})$$

と同値だが、これを‘1’(今の場合、死亡)の条件として用いることができる。

それによれば、両モデルとも、的中は115ケースで、的中率は $115/150 = 76.7\%$ である。

\hat{P}_i の推定値は表8.2の通りであるが、この場合はロジット・モデルのほうが当てはまりがよいようである。

表8.2 プロビット、ロジット・モデルによる投与量別死亡確率 \hat{P}_i

対数濃度 X_i	0	1	2	3	4
プロビット	0.139	0.294	0.500	0.706	0.861
ロジット	0.089	0.238	0.500	0.762	0.911
データ	0.067	0.267	0.500	0.767	0.900

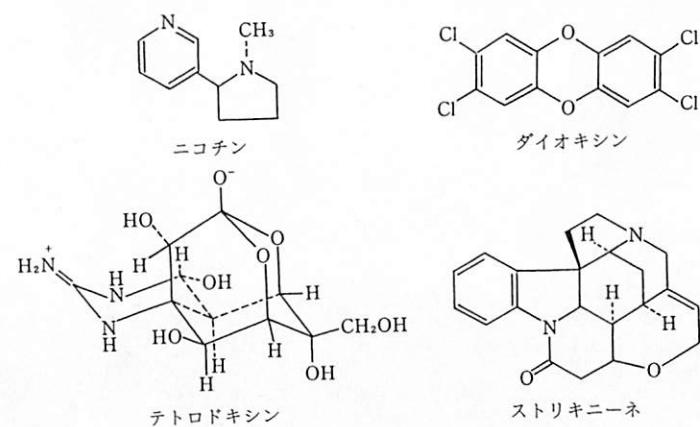
表8.1の量・反応関係に適用した、ロジット・モデルの推定が当てはまりがよい。これは出力結果の R^2 に多少とも表れているが、このことから、 R^2 の見方は注意を要すべきことがわかる。

※推理作家も利用する量・反応関係※

人体にとって有用などのような物質も大量に摂取すれば、人体にとって危険である。アルコール、食塩にしてしかりである。西洋医学史にその名を残す医学者にして自然学者、哲学者パラケルスス(P. A. Paracelsus, 1493-1541)も「すべての物質は毒である。物質にして毒でないものは何もない。毒と薬を分つのは量の適切な区別だけである」という言葉を残している。いろいろな物質の体重1kg当たりの半数致死量(LD 50)を挙げておこう。

物質名	LD 50(mg/kg)	備考(生活関連など)
エチルアルコール	10,000	
塩化ナトリウム(食塩)	4,000	
硫酸第1鉄	1,500	
モルヒネ(硫酸塩)	900	
フェノバルビタール(ナトリウム塩)	150	
ピクロトキシン	5	
ストリキニーネ(硫酸塩)	2	横溝正史の作品(推理小説)に登場
ニコチン	1	タバコ
d-ツボクラリン(クラーレ)	0.5	インディアンの矢毒。高木彬光の作品
ヘミコリニウム-3	0.2	
テトロドキシン	0.10	フグの卵巣、肝臓、内田康夫の作品
ダイオキシン(TCDD)	0.001	人工物質中最強の毒。ゴミ焼却で発生
ポツリヌス菌素	0.00001	過去に食品中毒の事例、猛毒

そのほか、コカイン(コカの葉、S. フロイトの発見)、アヘン(主要成分がモルヒネ)などの植物性アルカロイドの他、クロロフォルム、青酸、ヒ素などもよく知られている。



8.4 説明変数が2個以上の場合*

変数の個数を増やすことを考えよう。

例8.4 〈進路選択のデータ〉 表8.3はある大学において開設されているA, Bの2コースを30人の学生がどのように選択したかを示したデータである。

表8.1の量・反応関係の例では、説明変数がただ一つ X だったが、この場合は、学生のコースの選択を説明する変数は性別、前期試験の成績、専攻の三つである。プロピット、ロジット・モデルは回帰分析と同じく、表8.3のように Y_i が($p+1$)個の変数 $X_{1i}, X_{2i}, \dots, X_{pi}$ に依る場合に簡単に拡張できる。すなわち、(8.2)の Y_i^* が

$$(8.12) \quad Y_i^* = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

で与えられ、その符号によって Y_i が決定されるとする。8.1節のように、 F を $-\varepsilon_i$ の累積分布関数(プロピット・モデルの場合は標準正規分布、ロジット・モデルの場合はロジスティック分布)とすると、 $Y_i=1$ となる確率は

$$(8.13) \quad F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi})$$

で与えられる。

したがって、尤度関数は

$$(8.14) \quad L(\beta_0, \beta_1, \dots, \beta_p)$$

表8.3† 学生のA, Bの2コースの選択

学 生	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
コース選択 ^①	1	1	0	1	0	1	0	0	1	0	1	0	0	1	0
性 別 ^②	0	0	1	0	1	0	1	1	0	1	0	1	1	0	0
前期試験成績	89	85	84	83	81	88	87	86	86	84	87	83	82	83	82
専 攻 ^③	0	0	1	0	1	1	0	0	1	0	0	1	1	0	0

学 生	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
コース選択 ^①	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0
性 別 ^②	0	0	0	0	1	1	1	0	0	0	0	1	1	1	0
前期試験成績	86	78	77	85	75	80	82	79	83	82	89	86	80	81	78
専 攻 ^③	0	0	1	1	1	0	0	1	1	1	0	1	1	0	0

1) 0:Aコース, 1:Bコース, 2) 0:男, 1:女, 3) 0:文科系の専攻, 1:理科系の専攻

*) 多次元正規分布、若干の線形代数(ベクトルと行列)の知識を必要とするので、概論と適用例を与える。多次元正規分布については、本シリーズI『統計学入門』8.3節参照。

$$= \prod_{Y_i=0} F(\beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}) \times \prod_{Y_i=1} [1 - F(\beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi})]$$

であり、これを最大化することによって最尤推定量を求めることができる。いま、モデルをベクトル表示するため

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad X_i = \begin{pmatrix} 1 \\ X_{1i} \\ \vdots \\ X_{pi} \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

とし、 $\hat{\beta}$ を β の推定量である p 次元のベクトル、 $f(z)$ を F の密度関数とする、 $\hat{\beta}$ は一致推定量であり、その漸近分布は

$$(8.15) \quad \sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, A^{-1})$$

となる^{*)}。ここで A は

$$A = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \frac{f_i^2}{F_i(1-F_i)} X_i X_i'$$

$$F_i = F(X_i' \beta), \quad f_i = f(X_i' \beta)$$

で与えられる。 n が十分に大きければ、この A は、事実上

$$\hat{A} \doteq \frac{1}{n} \sum_{i=1}^n \frac{\hat{f}_i^2}{\hat{F}_i(1-\hat{F}_i)} X_i X_i'$$

$$\hat{F}_i = F(X_i' \hat{\beta}), \quad \hat{f}_i = f(X_i' \hat{\beta}) \quad (\text{ただし, } f=F')$$

によって求めることができる。この結果から、各 $\beta_0, \beta_1, \dots, \beta_p$ について、回帰分析と同じく t -検定を行うことができ、検定における自由度は $n-(p+1)$ である。

反応の確率 $P_i = P(Y_i=1|X_i)$ そのものは、(8.9) とおなじく

$$\hat{P}_i = F(X_i' \hat{\beta})$$

で推定され、その漸近分布は

$$\sqrt{n}(\hat{P}_i - P_i) \xrightarrow{d} N(0, f_i^2 \cdot X_i' A^{-1} X_i)$$

で与えられる。 A^{-1} の対角要素は推定量の分散、その他の要素は共分散であるから、 $p=1$ の場合この結果は(8.10) と同一となる。

*) 左辺の確率変数の確率分布が、 $n \rightarrow \infty$ のとき、右辺の確率分布となること(法則収束、分布の収束)をこのように簡略に記す。 $'d'$ は ' \mathcal{L} ' と記すこともある。

例 8.5 <モデルの推定：進路選択のデータ> 表 8.3 の進路選択のデータから、

X_1 ：性別、 X_2 ：前期試験の成績、 X_3 ：専攻、
としプロビット、ロジット・モデルによる分析を行ってみよう。その結果はプロビット・モデルでは

$$Y^* = -27.56 - 1.358X_1 + 0.337X_2 - 0.447X_3 + \varepsilon$$

ロジット・モデルでは

$$Y^* = -50.26 - 2.526X_1 + 0.612X_2 - 0.610X_3 + \varepsilon$$

となり、またこれらの正、負から、コース A, B の選択が決定される。

これから、 t 値を $t_1 = -1.358/0.634 = -2.134$ (プロビット)、 $t_1^* = -2.526/1.218 = -2.074$ (ロジット)などのように計算して、有意水準 5% で $H_0 : \beta_i = 0$ の検定を行えば、 β_1, β_2 では帰無仮説は棄却され、 β_3 では採択される。ただし、 β_1, β_2 では片側検定、 β_3 では両側検定を行った。すなわち、女性の方がより B コースを選択し、前期成績が良い程 A コースを選択する傾向があるが、専攻による選択には統計的な有意の差はないといえる。

モデルの検定 複数の説明変数があるとき、重回帰分析と同じように、いくつかの母数 β についての仮説を同時に検定したいことがある。たとえば、2つの薬剤 A, B を同時に与えた場合の害虫駆除の効果を調べるとしよう。 X_1, X_2 を薬剤 A, B のそれぞれの濃度とすると、帰無仮説は薬剤が死亡率に影響しないことであって

$$H_0 : \beta_1 = 0 \text{かつ } \beta_2 = 0$$

対立仮説は

$$H_1 : \beta_1 \neq 0 \text{ または } \beta_2 \neq 0$$

である。このように帰無仮説が複数の制約式からなるときは、 t -検定ではなく、 χ^2 分布に基づく尤度比検定を行う。

i) H_0 が正しいとしてモデル(上の例では X_1, X_2 を含まず、定数項だけのモデル)を推定し、その対数尤度 $\log L_0$ を計算する。なお、対数尤度はほとんどの統計パッケージのプログラムでは自動的に計算される。

ii) H_0 が誤りとしてすべての説明変数を含むモデルを推定し、対数尤度 $\log L_1$ を計算する。

iii) 帰無仮説に含まれる制約式の数を q とすると

$$(8.16) \quad T = -2(\log L_0 - \log L_1)$$

※ロジスティック曲線から「一般システム論」へ※

「ロジスティック」とはいったいどんな意味だろうか。‘logistic’を研究社『大英和辞典』で引いてみると、まず「論理学の」「数学的論理学の」という訳があり、‘logic’, ‘logical’と同語源とわかる。ギリシア語のロゴス logos(論理、理性、理法、言葉)である。さらに「算定(曲線)の」という訳語もあり、それで、人口の生(成)長の法則をあらわす曲線の名となったのである。たしかに、図 8.2 のロジスティック曲線には横軸に平行な漸近線があり、人口の飽和点を示している。

だが、かの有名なマルサス^{*}は「人口は制限せられなければ 25 年ごとに倍加する。いいかえれば、幾何級数的に増加する」と言った。しかし、生物学者ベルタランフィー(von Bertalanffy, 1901-1972)は、生物学の領域を超えて「一定の自然法則は経験的基礎によるだけではなく純形式的なやり方でも到達しうる」として、先駆的な一般システム理論^{**} General System Theory を提唱した。自然科学から社会・人文科学、哲学をも視野に入れる「システム理論」の始まりである。

「システム」として、一般に量 Q の生長が 5 割(瞬間)増加の微分方程式

$$(\#) \quad dQ/dt = 0.5Q$$

であらわされるとして、解として

$$Q = ce^{0.5t} = c(1.311)^t$$

が得られる。これは極端に速いが、増加率を 0.5 ではなく $\beta = (\log_e 2)/25$ をとると、速さはマルサスのものに一致する。

次に微分方程式(#)に、 Q が大きいときに増加に対して効く‘ブレーキ’の項——たとえば、資源の制約——を入れて、右辺を二次式

$$(\# \#) \quad dQ/dt = 0.5Q - 0.1Q^2$$

とすると、解は、今度はロジスティック曲線

$$Q = \frac{c \cdot 0.5e^{0.5t}}{1 + 0.1c \cdot e^{0.5t}}$$

となる。ブレーキが効いて、 $t \rightarrow \infty$ の長期で $Q \rightarrow 5$ となって生長が窮屈的に飽和に達する。

*) 「初版・人口の原理」高野・大門訳、岩波文庫、第 2 章

**) 「一般システム理論」長野・太田訳、みすず書房



ベルタランフィー
L. v. Bertalanffy
(1901-1972)

現代の生物学者。また、今日の科学方法論の重要な一翼を担う「システム理論」の創始者。「生命」「人間とロボット」などは多大な影響を与えた。

は、帰無仮説が正しい場合、漸近的に自由度 q の χ^2 分布 $\chi^2(q)$ に従うことが知られている。したがって、この T を検定統計量とし、 χ^2 分布によって仮説検定を行う。

例 8.6 <変数の影響：コース選択のデータ> 表 8.3 のコース選択データから性別と前期成績がコースの選択に影響するかどうか、帰無仮説 $H_0 : \beta_1 = 0$ かつ $\beta_2 = 0$ の検定を行おう。この場合、

$$\text{プロビット・モデルで, } \log L_0 = -19.47, \quad \log L_1 = -10.31,$$

$$\text{ロジット・モデルでは, } \log L_0 = -19.47, \quad \log L_1 = -10.19,$$

であり、 T の値はプロビットに対し $T = 18.32$ 、ロジットに対し $T = 18.56$ である。有意水準を 1% とすれば、自由度 2 の χ^2 分布から $\chi^2_{0.01}(2) = 9.2103$ であるから、プロビット、ロジット・モデルいずれでも H_0 は棄却され、性別と前期成績はコースの選択に影響していると考えられる。

このように、8.2 節以下で見たように、正規分布 $\phi(z)$ 、ロジスティック分布 $\Lambda(z)$ を用いて、幅広い応用性をもつ質的な二値データを分析するモデルを作成することができるのである。

<><><> 第8章 練習問題 <><>

8.1 <線形確率モデルとの比較> i) 線形確率モデルを用いて、表 8.1 の量・反応関係を記述する a_0, a_1 を求めよ。

ii) この推定値 \hat{a}_0, \hat{a}_1 を用いて、死亡確率 P_i を投与量別に推定せよ。

iii) 表 8.2 の投与量別推定死亡確率 \hat{P}_i をグラフにプロットし、ii)の結果を加えて書き入れよ。

iv) この線形確率モデルで、 $\hat{P}_i > 1$ または $\hat{P}_i < 0$ となる X_i の範囲を求めよ。

8.2 <LD 50 など> プロビット、ロジット・モデルで、 $Y_i = 1$ となる反応確率 $P = F(\beta_0 + \beta_1 X)$ ($F = \phi, \Lambda$)

に対し、次の X の値をそれぞれ求めよ。

i) $P = 1/2$ となる X の値

ii) $|dP/dX|$ が最大になる X の値とそれに対する $|dP/dX|$ の値

8.3 <両モデルの推定値比較> プロビット・モデルとロジット・モデルによる、 β_1, β_2 の推定値を比較して、対応する値の比が、 $\pi/\sqrt{3}$ 程度になることを、本章のいろいろなデータの分析結果から確認せよ。

8.4 <個別ケース予測> プロビット・モデル、ロジット・モデルによって表 8.3

のコース選択データを分析した結果を用いて、データの30人の学生のコース選択を予測せよ。実際のコース選択との一致は、30人中何例、何%となるか。

8.5[†] 〈市場調査〉 消費者24人を対象に、車を購入したかどうかのアンケートを実施したところ、次のような結果を得た。

番号	1	2	3	4	5	6	7	8	9	10	11	12
車購入の有無 ¹⁾	0	1	1	1	0	0	0	1	1	1	1	0
収入(万円)	300	500	450	600	400	450	500	400	700	600	550	600
性別 ²⁾	1	0	0	1	0	1	1	0	0	0	1	1
居住地域 ³⁾	1	0	1	0	1	0	1	0	0	1	1	1

番号	13	14	15	16	17	18	19	20	21	22	23	24
車購入の有無 ¹⁾	1	1	0	1	1	0	0	1	0	1	1	0
収入(万円)	600	650	400	500	550	650	600	450	400	700	400	500
性別 ²⁾	0	1	0	0	1	0	1	0	1	0	0	1
居住地域 ³⁾	1	0	1	0	1	1	0	1	0	1	1	0

1) 0: 購入せず 1: 購入 2) 0: 男性, 1: 女性 3) 0: 都市に居住, 1: 農村に居住

- i) 車購入を予測するプロピット、ロジット・モデルを作り、それを推定せよ。
- ii) 適当な帰無仮説(および対立仮説)を設定し検定せよ。
 - a) 収入は車の購入に影響するか。
 - b) 性別と居住地域は(同時に)車の購入に影響するか。
- iii) 収入500万円の都市に居住する男性が車を購入する確率を求めよ。また、その確率の95%信頼区間を求めよ。

第9章 ベイズ決定

ベイズ統計学は、ベイズの定理のもつ多くの有用な性質を活用する統計学である。この統計学によれば、人間のもつ予想、確信、信念なども一種の‘データ’であり、「主観確率」として判断や決定のためにとり入れることができる。われわれがしばしば経験する情報不十分という状況の中で、この統計学が有効性を発揮するわけもあり、ふつうの統計的推測の理論(ネイマン・ピアソン理論)とは違った新しさがある。

この章は、ベイズ統計学のエッセンスを簡潔に述べた、独立した内容をもつ章である。

9.1 ベイズの定理

ベイズ決定とは、ベイズの定理を用いた統計的決定である。一般に、ベイズの定理を積極的にデータを扱う考え方の中に利用する統計学の理論を、ベイズ統計学 Bayesian statistics といい、近年急速に普及している。ベイズ決定はその中の1つの分野である。

ベイズの定理は、原因の確率に関する定理である。いま、 k 個の可能なすべての「原因」を H_1, H_2, \dots, H_k とする。異なる原因 H_i, H_j は同時には成立せず、互いに排反($H_i \cap H_j = \emptyset$)としておく。各 H_1, H_2, \dots, H_k の確率は、最初は $P(H_1), P(H_2), \dots, P(H_k)$ とする。この確率は、あらかじめ予想、確信に対して、数学的に簡単な数字を与えたもので、事前確率 prior probability と呼ばれる。この事前確率こそ、ベイズ的な(ベイジアン)考え方の基であり、その特色である。

いま、事象 A を考えこれを結果とする。各原因からこの A が生じる条件付確率は、 $P(A|H_1), P(A|H_2), \dots, P(A|H_k)$ であるとしよう。ベイズの定理は、 A を生じさせた各原因の可能性は、条件付確率として

$$(9.1) \quad P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{\sum_j P(H_j) \cdot P(A|H_j)}$$

によって計算されるというものである。この $P(H_i|A)$ を $P(H_i)$ に対して、事後確率 posterior probability という。事後確率には結果 A が含まれているから、これをデータ data と読みかえれば、「ベイズの定理に基づく(利用する)統計学」が成立する可能性があることがわかる。

例 9.1 〈ベイズの定理の計算と実験〉 外見からはまったく見分けのつかない 2 つのつぼがあり、1 つのつぼには、70 個の赤玉と 30 個の青玉が入っている、もう 1 つのつぼには、30 個の赤玉と 70 個の青玉が入っている。赤玉の多い方のつぼを H_R 、青玉の多いつぼを H_B とする(あるいは、赤玉の割合(確率)を θ として、 $\theta=0.7$ の方と $\theta=0.3$ の方といつてもよい)。いま、ランダムにつぼを選ぶとする。すなわち、 $P(H_R)=1/2$ 、 $P(H_B)=1/2$ 。ランダムだから、他に理由がなければ等確率である。これを、理由不十分の原則 principle of insufficient reason という。

いま、選ばれたつぼから玉を 1 つずつとり出して、赤か青かを記録し、1 回ごとにもとへ戻すものとする(復元抽出)。この方法で 12 個の玉をとり出したとき、赤玉 8 個、青玉 4 個を得た。このデータ(D)を元にして、どちらのつぼから選ばれたかが計算できるだろうか。

このとき、 $P(H_R|D) > 1/2$ となるのは当然であろう。この事後確率をベイズの定理(9.1)から計算すると、母数 $\theta=0.7$ および 0.3 の二項分布を用いて

$$\begin{aligned} P(H_R|D) &= \frac{P(H_R) \cdot P(D|H_R)}{P(H_R) \cdot P(D|H_R) + P(H_B) \cdot P(D|H_B)} \\ &= \frac{(1/2) \cdot {}_{12}C_8 (0.7)^8 (0.3)^4}{(1/2) \cdot {}_{12}C_8 (0.7)^8 (0.3)^4 + (1/2) \cdot {}_{12}C_4 (0.7)^4 (0.3)^8} \\ &= 0.97 \end{aligned}$$

と計算できる。この確率は意外に高い。実験(エドワーズ)によると、多くの人々はデータに基づいて 0.75 前後の値と想像するようである。ベイズの定理は計算規則に基づいて一応の基準的な値を与える。

ベイズの定理を統計学にとり入れるために、例 9.1 でも見たように「原

因」として母数 θ をとり、「結果」として標本のデータ z をとり^{*}、さらに事前確率 $P(H_i)$ を $w(\theta_i)$ 、事後確率 $P(H_i|A)$ を $w'(\theta_i|z)$ 、さらに、結果の確率分布 $P(A|H_i)$ を $f(z|\theta_i)$ と表す^{**}と、 θ の事後確率

$$(9.2a) \quad w'(\theta_i|z) = \frac{w(\theta_i) \cdot f(z|\theta_i)}{\sum_j w(\theta_j) \cdot f(z|\theta_j)}$$

を得る。 θ が連続的なら、和を積分にかえ

$$(9.2b) \quad w'(\theta|z) = \frac{w(\theta) f(z|\theta)}{\int_\theta w(\theta) f(z|\theta) d\theta}$$

となる。ここに θ は、可能な母数の値の集合で、母数空間 parameter space と呼ばれる。これら $w(\theta)$ 、 $w'(\theta|z)$ を θ の確率分布として見て、事前確率分布、事後確率分布という。

(9.2a)、(9.2b) は、 w 、 w' および尤度関数 $f(z|\theta)$ の関係として、簡単に

$$(9.3) \quad w'(\theta_i|z) \propto w(\theta_i) \cdot f(z|\theta_i)$$

と書くことができる。データ z が尤度(そこでは、データは固定され、 θ が動く)という形で入っていることが重要である。分母が定数として省略されている。分母は、全確率が 1、すなわち規格化の条件

$$(9.4) \quad \sum_i w'(\theta_i|z) = 1, \quad \int w'(\theta|z) d\theta = 1$$

を満たすためのものであるから、(9.3) では明らかなものとして省略してもよいのである。いずれにせよ、重要なことは、母数 θ は確率分布(9.2)、(9.3) をもつ確率変数であるという考え方である。

9.2 事前確率分布と事後確率分布

「事前確率」は情報が十分にないときに役に立つ考え方である。

例 9.2 〈新薬開発〉 開発中の新薬が偽薬(有効成分を含まない比較対照薬、プラセボ placebo という)と比較して、有効となる確率を θ ($0 < \theta < 1$) と

^{*}) 本来は確率変数であるがデータとして与えられているので小文字 z を用いるのが普通である。

^{**}) θ に対する z の確率分布であるから $f_\theta(z)$ と書くのが本シリーズの原則だが、ベイズ統計学では条件付の確率分布と考えて本記法を採用する。

する。もし、 θ の値が十分大であれば、新薬開発をさらに進めてよい。したがって、 $0 < \theta < 1$ のどのような値となるかが、重要な問題となる。そこで、 θ に確率分布 $w(\theta)$ をこれから考える。

※学者の無節操——ラプラスの場合※

ベイズの定理の発見者トマス・ベイズ(Thomas Bayes, 1702(?)~1761)は、この簡単で基本的な大定理を導いた人としては有名である。彼自身は、本職は終生牧師であった。この定理以外に、スターリングの公式に関する論文がある他は、「学者」として知られている人ではない。現に、生年月日も確かではなく、ベイズの定理(の原型)も、生命保険の考え方の創始者の一人として有名なプライス(Richard Price, 1723~1791)が、後日友人として世に紹介したものである。

ベイズの定理を、今日のように、「原因の確率」probabilité des causesとして仕上げたのは、いうまでもなく、古典確率論の完成者にして近代確率論の創始者、大著『確率の解析的理論』Theorie analytique des Probabilitésの著者であるラプラス(Simon de Laplace, 1741~1821)である。

しかし、このラプラスは、ベイズの業績を用いるに際して、ベイズ(の名)を引用していない^{*)}。ラプラスは、「世渡り」のうまい、無節操な人であったようである。ちなみにラプラスのこの大著の第1版にある、ナポレオンへの献呈の辞は話題を呼んだ。

ナポレオン大帝へ捧ぐ。陛下、臣の天体力学論を、献上申し上げました
節にお示し下さいました陛下の御好意に対しまして、臣は確率計算に関する著作をも陛下に献呈申し上げたいという願望をおさえきれないでございます。……陛下のいとも下賤なる、いとも従順なる下僕、忠良なる臣、ラプラスより

『確率の解析的理論』のナポレオンへの献辞

ナポレオン没落後王政復古の世となったとき、ラプラスは第2版以降この献呈の辞を全文削除した。この無節操が後世、世人の非難的になったという。ナポレオンが陸軍士官学校時代、教官としてその試験をしたのがラプラスであったのだが…。

^{*)} Encyclopedia of Statistics. ただし、有名なトドハンター『確率論史』の邦訳註によると、これは厳密には事実と相違するという。

比較対照の試行で、新薬がプラセボに比して効くと判定されることを「成功」とすれば、 n 回の試行中 x 回の成功を得る確率は、二項分布

$$(9.5) \quad f(x|\theta) = {}_nC_x \theta^x (1-\theta)^{n-x}, \quad x=0, 1, \dots, n$$

で二項分布 $Bi(n, \theta)$ となる。 θ の事前確率分布 $w(\theta)$ さえうまく定めれば、前節で見たように、事後確率分布をベイズの定理から計算できる。(9.2b)あるいは(9.3)を用いればよいのである。この際、一般に、事前確率分布は、 θ の実験前の予想をなるべく正確に反映するものであると同時に、(9.2b)の計算が容易になるように定めなければならない。

二項分布の形の尤度(9.5)に対しては、その式の形に注目して θ の事前確率分布としてベータ分布 $Be(\alpha, \beta)$

$$(9.6) \quad w(\theta) = \theta^{\alpha-1} (1-\theta)^{\beta-1} / B(\alpha, \beta), \quad 0 < \theta < 1$$

ただし

$$B(s, t) = \int_0^1 u^{s-1} (1-u)^{t-1} du \quad (\text{ベータ関数})$$

を採用すると、(9.3)はごく自明な計算となる。つまり

$$(9.7) \quad w'(\theta|x) \propto \theta^{\alpha+x-1} (1-\theta)^{\beta+(n-x)-1}$$

であって、 θ の事後確率分布は $Be(\alpha+x, \beta+(n-x))$ となる。(9.7)の比例定数が $1/B(\alpha+x, \beta+n-x)$ であることはいうまでもない。結局、この場合、同じベータ分布の中で

$$\alpha \rightarrow \alpha+x, \quad \beta \rightarrow \beta+(n-x)$$

の変換をするだけでよい。

一般に、尤度 $f(x|\theta)$ に対して、 $w(\cdot)$ と $w'(\cdot|x)$ が同一種類の分布(分布族)に属するならば、標本値 x は単に同一分布族内の変換を引きおこすだけである。このとき、 w , w' の分布族を f の自然な共役分布の族 natural conjugate family という。二項分布 $Bi(n, \theta)$ の自然な共役分布は、ベータ分布 $Be(\alpha, \beta)$ である。自然共役分布族をうまく——やや直観的に——見出すことが、ベイズ統計学の極意である。

例9.3 〈続〉 新薬開発の中間過程では、 θ の値は当面まったく手探りでわからないとすれば、事前確率分布は $Be(1, 1)$ 、つまり一様分布

$$(9.8) \quad w(\theta) \equiv 1 \quad (0 < \theta < 1)$$

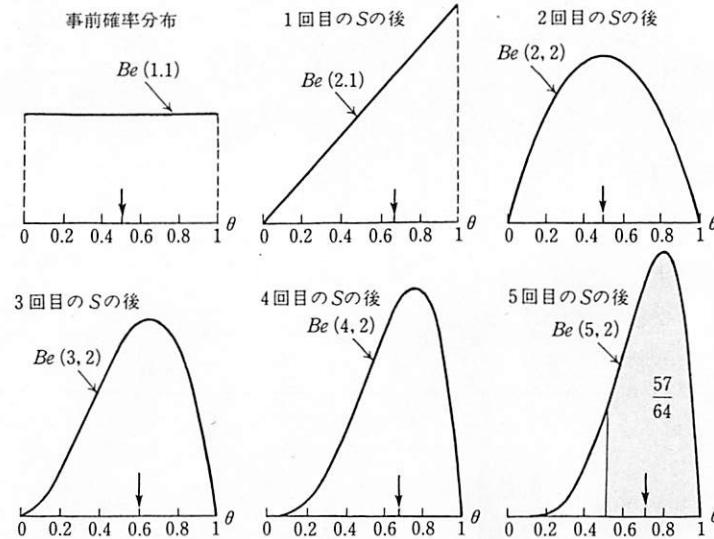


図 9.1' 事後確率分布の変化(新薬開発の例)

分布が各段階で更新され、 θ の大きい方(新薬にとって有利な方)へ集中していく様子がわかる。分布はベータ分布で、矢印はその期待値(平均)の位置である。

であるとするのが自然である^{*)}。

次に、 n 回の実験(試行)中の成功(S)、失敗(F)の回数を $x, n-x$ とすると、この時点での事後確率分布は、(9.7)から $Be(1+x, 1+n-x)$ となる。 $n=1, 2, \dots$ と順次見ていくとき、結果は S が多く

S, F, S, S, S, \dots

であったとする。この系列からは、新薬がプラセボに比して効くという確率は、徐々に高まってゆくはずである。実際、事後確率分布は、事前確率分布 $Be(1, 1)$ から、順次 $Be(2, 1), Be(2, 2), Be(3, 2), Be(4, 2), Be(5, 2), \dots$ と変ってゆくが(図 9.1)，確率分布の形から、新薬が有効である確率 θ が次第に 1 の周囲に寄っていくことがわかる。 $n=5$ では、 $x=4, n-x=1$ で、事後確率分布は $Be(5, 2)$ である。この確率分布においては、新薬にとって有利になる。つまり $\theta \geq 1/2$ となる確率は、事後確率分布で計って

^{*)} 以下、巻末 Berry による。

$$(9.9) \quad P(\theta \geq 1/2) = \int_{1/2}^1 \theta^4 (1-\theta) d\theta / B(5, 2) \\ = 57/64$$

であって、かなり確からしいこととなる。

ベータ関数 $B(m, n)$ の計算は、 m, n が正整数なら

$$B(m, n) = (m-1)! (n-1)! / (m+n-1)!$$

となることを知っておくとよい。

また、確率変数 θ の期待値 $E(\theta)$ も、ベータ分布 $Be(\alpha, \beta)$ の期待値

$$(9.10) \quad \mu = \alpha / (\alpha + \beta)$$

から求められる。この場合も図 9.1 に見るように、次第に 1 に近づいている。

なお、事前確率分布 $Be(1, 1)$ から求められた、各時点の事後確率の列

$$\dots, Be(2, 1), Be(2, 2), Be(3, 2), Be(4, 2), Be(5, 2), \dots$$

相互の間に、「事前」「事後」の関係がある。すなわち、(9.6)で事前確率分布として、 $\alpha=2, \beta=1$ とし、さらに $n=1$ の確率データをとって、 F が生じたとする。 $x=0, n-x=1$ であるから、事後確率分布として(9.7)で $Be(2, 2)$ が得られる。これを $Be(2, 1) \xrightarrow{F} Be(2, 2)$ と書くと、 S, F, S, S, S, \dots の列は、この事前、事後の組み合わせによって次々と

$$Be(2, 1) \xrightarrow{S} Be(2, 2) \xrightarrow{F} Be(3, 2) \xrightarrow{S} Be(4, 2) \xrightarrow{S} Be(5, 2) \Rightarrow \dots$$

のように、事後確率分布を生み出す。

一般に第 t 期の事後確率分布が、第 $(t+1)$ 期においては事前確率分布として作用し、第 $(t+1)$ 期の標本データによって新たな事後確率分布が生れることが(9.3)から証明できるが、この性質を、ベイズの定理によるベイズ更新 Bayesian updating という。これはベイズの定理の有する有用な一般的性質の一つである。

自然な共役分布でよく知られるのは

a) 正規分布 $N(\theta, \sigma^2)$ の平均の母数 θ について、正規分布 $N(\lambda, \tau^2)$

b) ポアソン分布 $P_o(\theta)$ に対して、ガンマ分布 $Ga(\alpha, \lambda)$

がそれぞれ自然共役分布となる場合である。とくに a) はしばしば用いられる。

$z=(x_1, x_2, \dots, x_n)$ を正規母集団 $N(\theta, \sigma^2)$ からの、大きさ n の標本とするならば、尤度は

$$(9.11) \quad f(z|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\theta} \exp\left\{-\frac{(x_i-\theta)^2}{2\sigma^2}\right\}$$

$$= \exp\left\{-\frac{(\bar{x}-\theta)^2}{2\sigma^2/n}\right\} \cdot \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{\sum(x_i-\bar{x})^2}{2\sigma^2}\right\},$$

$$(\bar{x} = \sum x_i/n)$$

となり、 θ の関数は第一因数だけとなる。実際、(9.2b)では、第二因数は消える。これから、第一因数中の統計量 \bar{x} を知れば、 θ を知るために十分 sufficient である。(9.3) は

$$(9.12) \quad w'(\theta|z) \propto w(\theta) \cdot \exp\left\{-\frac{(\bar{x}-\theta)^2}{2\sigma^2}\right\}$$

となる。この式を見れば、 $w(\theta)$ も正規分布であって

$$(9.13) \quad w(\theta) = \frac{1}{\sqrt{2\pi}\tau} \exp\left\{-\frac{(\theta-\lambda)^2}{2\tau^2}\right\}$$

とすれば、計算上好都合であることがわかる。(9.12)の積を実行すれば、 \exp の中の二次式の操作だけで

$$(9.14) \quad w'(\theta|z) \propto \exp\left\{-\frac{(\theta-\lambda')^2}{2\tau'^2}\right\}$$

となる。ここで

$$(9.15a) \quad 1/\tau'^2 = (1/\tau^2) + (n/\sigma^2)$$

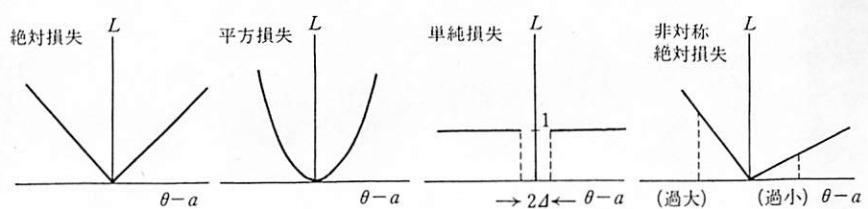
$$(9.15b) \quad \lambda' = \frac{(1/\tau^2)\lambda + (n/\sigma^2)\bar{x}}{(1/\tau^2) + (n/\sigma^2)}$$

である。すなわち、事後確率分布は、事前確率分布の平均 λ と、標本平均 \bar{x} の比 $(1/\tau^2) : (n/\sigma^2)$ の内分点 λ' を新たな平均、 $(1/\tau^2)$ と (n/σ^2) の和の逆数を新たな分散とする正規分布となる。とくに、標本の大きさ n が大きくなるほど、 λ' においては、標本平均 \bar{x} が大きく効き、 λ' は \bar{x} に引きつけられる。

(9.15)において分散の逆数 $h = 1/\tau^2$, $h_s = 1/\sigma^2$ が表れている。これを精度 precision という。精度は、事前分布のそれと標本のそれとの和になっている。

9.3 ベイズ推定

一般に、事後確率分布は、母数 θ の確率分布を示すことにより、 θ のありか



最も通常の場合で、単 ‘はずれ’ が大きくなるほど、制裁の程度が加速度される。
‘はずれ’ (Δ) 以内の誤差は ‘当り’ と見なす) の見ている。

図 9.2 推定のためのいろいろな損失関数
はずれた場合の考え方(評価)に、いろいろなものがある。なお、この図は θ も a も連続的な場合を示してある。

の予想を示したものである(たとえば、図 9.1)。しかし、それは一つの値を示したものではない。事後確率分布はそれぞれの値が出る可能性を示しているが、どの値が出るかを示してはいない。

θ の値として一つの値(a としよう)を示すことは一つの決定であり、統計学的には推定 estimation の問題となる。(9.2)あるいは(9.3)によって与えられる事後確率分布 $w'(\theta|z)$ から、 θ を推定する a の値を、 z の関数(推定値)としてとり出すには、どのようにすればよいのであろうか。図 9.1 で示したような、事後確率分布の期待値(平均値)は、一つの基準であるが、それだけではない。

一般に a と θ の(真の値)はなるべく近い方がよく、離れるに従い望ましくないことになる。そこでこの推定のペナルティー(これを損失関数 loss function という)を $L(\theta, a)$ とおき、これを基準に考えよう(図 9.2 参照)。

期待損失の最小化の原理 このときまでは、 $L(\theta, a)$ の期待値(平均値)

$$(9.16) \quad \sum_z f(z|\theta) \cdot L(\theta, a(z))$$

または

$$\int f(z|\theta) \cdot L(\theta, a(z)) dz$$

を最小にする z の関数 $a(z)$ として、最適な行動を求める。

これらの量は、まだ θ の関数であり、もとより θ は未知であるから、このままで最小化は実行できない。 θ はもともと $w(\theta_i)$ という確率分布(事前確率分布)をもっているのであるから、さらに期待値をとった

$$(9.17) \quad \sum_i w(\theta_i) \{ \sum_z f(z|\theta_i) \cdot L(\theta_i, a(z)) \}$$

が、推定の評価基準となる(θ が θ_i でなく、連続的な値をとるときは、 \sum_i は積分となる)。(9.17)の和を組み替えベイズの定理(9.3)を用いれば、(9.17)の最小化は、 $\dot{z} \cdot \dot{z}$ ごとに

$$(9.18) \quad \sum_i w'(\theta_i|z) \cdot L(\theta_i, a(z))$$

を最小化することに帰する。すなわち、事後確率分布による損失関数の期待値(期待損失 expected loss)を最小化すればよいことがわかる。

この最小事後期待損失の基準は、一般的に9.5節の統計的決定問題の解によって示されるので、ここでは結果だけを用いる。

ベイズ推定 事後確率分布による損失 $L(\theta, a)$ の期待値(平均値)を最小にする $a(z)$ を実際に求めてみよう。以下では典型的な場合のみを扱う(図9.2)。

i) 絶対損失 $L(\theta, a) = |\theta - a|$ 一般に確率変数 X に対する最小化

$$(9.19) \quad \min_a E|X - a|$$

の解は、 $a = (X \text{ のメディアン})$ であるから

$$(9.20) \quad a(z) = (\text{事後確率分布 } w'(\theta|z) \text{ のメディアン})$$

となる。

ii) 平方損失 $L(\theta, a) = (\theta - a)^2$ 一般に確率変数 X に対する最小化

$$(9.21) \quad \min_a E(X - a)^2$$

の解は、 $a = E(X)$ であるから

$$(9.22) \quad a(z) = (\text{事後確率分布 } w'(\theta|z) \text{ の期待値})$$

となる。これが、もっともよく用いられる場合である。

iii) 単純損失 $\Delta > 0$ が十分小さい定数で

$$(9.23) \quad L(\theta, a) = \begin{cases} 0, & |\theta - a| \leq \Delta \text{ のとき} \\ 1, & |\theta - a| > \Delta \text{ のとき} \end{cases}$$

この場合には、 a を中心とする $\pm \Delta$ の区間 $[a - \Delta, a + \Delta]$ に最大の事後確率が入ればよいから

$$(9.24) \quad a(z) = (\text{事後確率分布 } w'(\theta|z) \text{ のモード})$$

となる。

iv) 非対称絶対損失、つまり i)の場合を非対称にして

$$(9.25) \quad L(\theta, a) = \begin{cases} k_0|\theta - a|, & a \geq \theta \\ k_1|\theta - a|, & a < \theta \end{cases}$$

過大推定と過小推定を分ける場合に相当する(ただし、 $k_0, k_1 > 0$)。途中を省略すると、結果は、 $c = k_1/(k_0 + k_1)$ として、

$$(9.26) \quad a(z) = (\text{事後確率分布 } w'(\theta|z) \text{ の } 100c \text{ パーセント点})$$

以上が、ベイズの定理を用いたベイズ推定の一般論である。

二項分布の θ のベイズ推定 尤度が二項分布(9.5)で、事前確率分布、事後確率分布がペータ分布(9.6), (9.7)のときには、標本データ z (今のは成功の回数 x)にもとづくベイズ推定 $a(z)$ (これを一般に $\hat{\theta}$ と表す)は、ii)の平方損失の場合、(9.10)を用いて

$$(9.27) \quad \hat{\theta} = \frac{\alpha + x}{\alpha + \beta + n}$$

であり、iii)の単純損失の場合、 $Be(\alpha, \beta)$ のモードが

$$(9.28) \quad M = \frac{\alpha - 1}{\alpha + \beta - 2}$$

であるから

$$(9.29) \quad \hat{\theta} = \frac{\alpha + x - 1}{\alpha + \beta + n - 2}$$

となる。

例 9.4 〈続〉 新薬のプラセボに対する有効性が、先のように、成功(S), 失敗(F)の列として

$$S, F, S, S, S, \dots$$

であったとする。事前確率分布を(9.8)のように一様分布 $Be(1, 1)$ としたとき、ベイズ推定は、 $n = 1, 2, 3, \dots$ となるに従い、平方損失の場合は

$$\hat{\theta}_1 = \frac{2}{3}, \quad \hat{\theta}_2 = \frac{2}{4}, \quad \hat{\theta}_3 = \frac{3}{5}, \quad \hat{\theta}_4 = \frac{4}{6}, \quad \hat{\theta}_5 = \frac{5}{7}, \dots$$

となり、絶対損失の場合は

$$\hat{\theta}_1 = 1, \quad \hat{\theta}_2 = \frac{1}{2}, \quad \hat{\theta}_3 = \frac{2}{3}, \quad \hat{\theta}_4 = \frac{3}{4}, \quad \hat{\theta}_5 = \frac{4}{5}, \dots$$

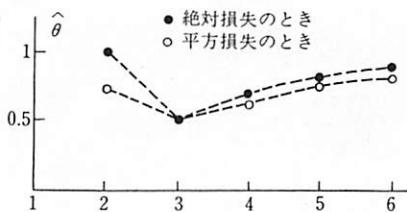


図 9.3 ベイズ推定(二項分布の場合)

S, F, S, S, S, \dots に基づくベイズ推定 $\hat{\theta}$ の図。事前確率分布は一様分布である。

となる(図 9.1, 9.3 参照)。時間(情報)とともに、 $\hat{\theta}$ がそれを反映して逐次的にベイズ更新 update される点に注目しよう。

正規分布の平均 θ のベイズ推定 正規母集団 $N(\theta, \sigma^2)$ からの標本データ $z = (x_1, \dots, x_n)$ にもとづいて、それぞれ、絶対、平方、単純の各損失を用いるベイズ推定 $\hat{\theta}$ は、正規分布のメディアン、平均、モードはすべて一致することから、(9.15b)によって

$$(9.30) \quad \hat{\theta} = \frac{(1/\tau^2)\lambda + (n/\sigma^2)\bar{x}}{(1/\tau^2) + (n/\sigma^2)}$$

で与えられる。ここに、 λ は事前確率分布(正規分布)の平均、 τ^2 は同じく分散、 σ^2 は母分散、 \bar{x} は確率平均 $\bar{x} = (x_1 + x_2 + \dots + x_n)/n$ である。

ここでは、 n の値は固定してある。これを固定標本数 fixed sample size の場合という。しかし、前節のベータ分布の例では、観測値の個数を 1 つずつ時間的に加えながら、 $n=1, 2, \dots$ のように、標本をとってゆく場合を考えた。そこで、いま $t=0, 1, 2, \dots$ 、時点を表そう。そして(9.15b)で、 $n=t$ 、 λ を $\hat{\theta}_t$ 、 τ^2 を $\hat{\theta}_{t+1}$ 、 $x=x_{t+1}$ とおく。この記号を用いて、(9.30)を書きなおすと

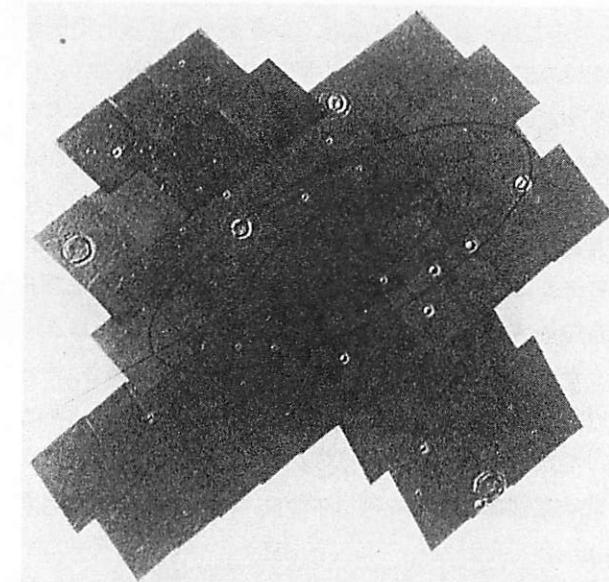
$$(9.31) \quad \begin{aligned} \hat{\theta}_{t+1} &= \frac{(1/\tau^2)\hat{\theta}_t + (1/\sigma^2)x_{t+1}}{(1/\tau^2) + (1/\sigma^2)} \\ &= \frac{\sigma^2\hat{\theta}_t + \tau^2x_{t+1}}{\sigma^2 + \tau^2} \quad (t=0, 1, 2, \dots) \end{aligned}$$

となる。

これは、1 つずつ観測される標本データ x_1, x_2, x_3, \dots によって、逐次、 θ を推定する推定量 $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \dots$ を作る漸化式となっている。これは制御理論で、カルマン・フィルター Kalman filter といわれる推定法のもっとも中心的な部分である。

※カルマン・フィルターもスペース・エイジを導いた※

人工衛星の制御にコンピューターが不可欠であることは、従来から常識である。むしろ、コンピューターによって人口衛星なるものが可能になったといっても、それほど言いすぎではない。具体的に用いられるソフトウェアは、位置、速度、姿勢の統計的推定法である。それは刻々入ってくる入力情報を次々とリアル・タイムで処理するものでなくてはならない。カルマン・フィルターは、その典型的なものである。これを用いると、目的天体上のある領域(ふつうは、橢円)に入る確率が、たとえば 50% になるように定めることができ、これによって、着地の精確さを保証することができる。この領域を、半数可達域という。



写真是、1976 年、NASA の「ヴァイキング 1 号」が、火星のクライス・プラニシア海へ着地する際の、半数可達域(内側)である。外側は 95% 可達域であり、これらは統計学的には同時信頼域である。

9.4 統計的決定理論

推定と検定の一般化 推定や検定の統計的推測には、どの方法で行うかという問題のあとにそれによった場合の判断(ないしは決定)の正しさという一般的な問題がある。

たとえば、新薬と通常薬を比較対照する実験で、ある検定法によれば差がないのに、差があると判定する確率(第1種の過誤の確率)は0.05を超えないが、差があるのに差がないと判定する確率(第2種の過誤の確率)は0.12であった。いま、第1種の過誤のコストを c_1 、第2種の過誤のコストを c_2 とすると、もし、正しい判断のコスト=0ならば、この検定法の平均的コストは

$$c = 0.05c_1 + 0.12c_2$$

となるであろう。この基準の c が小さい検定法が他にあれば、その方が‘正しい’検定法であることはいうまでもない。もちろん、コストの関数が基準であるとしてあるが、他の関数であっても同様のことがいえる。

推定の問題でも、推定がはずれた場合のコスト(損失)がいろいろの関数で(図9.2参照)表されている。どれを用いるかは時と場合によるが、どれが採用されたにせよ、対応する最適な推定量が(9.19)から(9.26)のように求められる。

ここで考え方を発展させて、ある基準で見た最も正しい判断を得ることを考えよう。そう考えると、問題や場合、状況にふさわしい基準を採用することにより、その場合の最適な判断を決めることができる。

これらのことから、推定も検定も、ある基準になる‘コスト’を最小にする統計量(推定なら推定量、検定なら検定統計量)を求めており、という一般的な見方ができる。統計的推測全体を一つの見地から統一的にとらえるこの考え方を、ワルドの統計的決定理論 statistical decision theory という。また、ここでコ

表9.1 仮説検定における2種類の過誤

	(採択) a_1	(棄却) a_2		(採択) a_1	(棄却) a_2
(仮説) θ_1	0	1	(仮説) θ_1	0	3
(対立仮説) θ_2	2	0	(対立仮説) θ_2	1	0

誤りに2種類あることはよく知られているが、ここで示すのはその重要さにはウェイトを付けることができるということである。

ストを表す関数が、一般の損失関数である。

推定の損失関数の例はすでに図9.2で挙げた通りであるが、検定ではどのような形になるであろうか。仮説を $\theta=\theta_1$ 、対立仮説を $\theta=\theta_2$ とすると、たとえば表9.1のようになる。仮説を「採択」、「棄却」することを a_1 , a_2 と略記した。左の場合は、 $c_1=1$, $c_2=2$ 、右の場合は、 $c_1=3$, $c_2=1$ で、左の場合は第1種の過誤をより重く見ていることがわかる。

推定にしても検定にしても、母集団の未知の母数 θ に対する、ある判断の行動 a の正しさの評価が問題になっている。推定では a は推定値(を与えるという行動)、検定では a は採択、棄却の行動である。損失関数は θ と a で決まるので、 $L(\theta, a)$ と書かれる。統計的決定理論は、損失関数 $L(\theta, a)$ を基準にして、最適な行動 a を、データ z を元にして(‘統計的’の意味)決定するという一般的な考え方のことである。このことを再確認するためには、(9.20), (9.22), (9.24), (9.26)などを参考にされたい。

例9.5 <ネルソン氏の雨の問題> 統計的決定理論の考え方は検定や推定の一般化だけあって、きわめて自然なものであり、広く日常生活の中にも見出すことができる。(以下は、巻末チャーノフ、モーゼスの挙げた例である)。

朝、家を出るとき、その日一日の天気の状態、晴、雨、曇、…が気になるものである。状態(θ)に関係して、何も用意しないで出るか、かさだけ持つか、レイン・コート、レイン・シューズを着用して‘完全武装’にするか、…いろいろと考える。その日の天気が完全予測できない以上、ただちにこれらの行動(a)が決まるとは考えられない。

そのために、ネルソン氏は毎朝、晴雨計(気圧計)の示度を読む。現代であれば、新聞の確率予報でもよい。そこには‘高’‘中’‘低’のどれか(z)が示されている。これは、統計学の‘データ’である。その日の天気が晴なら、‘高’を示す確率は高いであろう。したがって、このデータは確実とはいえないが、行動を決める基礎にはなる。

したがって、‘高’‘中’‘低’に従い、どの行動をとるか、すなわち、 z に対して a を対応させるルールはどのようなルールが最適 optimal か、という問題を考えることになる。これらは、一般に、不確実性下の意思決定 decision making under uncertainty といわれる。状態は

表9.2 ネルソン氏の状況

θ	a	(晴天用)	(レイン・コート)	(レイン・シューズ, かさ)
θ	z	(高)	(中)	(低)
a_1	z_1	0	1	3
a_2	z_2	0.60	0.25	0.15
a_3	z_3	5	3	2

5は'ずぶぬれ'になるコスト、右上の3は、'わざらわしさ'、'かっこ悪さ'のコスト、と考えられる。2は、正しい決定でも、雨という結果の悪さは避けられないことを意味するものとする。なお、「ネルソン」とは架空の人物名である。

a_1 : 晴天, a_2 : 雨天

行動は

a_1 : 晴天用服装, a_2 : レイン・コート,
 a_3 : レイン・コート, レイン・シューズ, かさ

データは、晴雨計の読みで

z_1 : 高(晴天), z_2 : 中(不明), z_3 : 低(雨天)

である。損失関数、データの確率分布の表は、表9.2のようになっている。

この2表のうち、右側のデータの確率分布の表が、母数 θ に対してデータの出方を示すという意味で、ふつう「統計的」といわれる部分である。問題は、晴雨計の読み z_1, z_2, z_3 に対して、それぞれどの行動 $a(a_1, a_2, a_3)$ が最適に対応するかである。この解は例9.7で与えられている。

統計的決定問題 統計的決定理論は、最適の行動 a を、データの関数

$$a=a(z)$$

として定めることを目的する。これを統計的決定問題 statistical decision problem という。この一般的な解法はベイズ推定に例をとて用いた((9.18) 参照)ように、最小事後期待損失 minimum posterior expected loss の基準である。

例9.6 <新薬開発> 新薬のプラセボに対する優位性(プラセボより効く確率) θ を、平方損失 $|\theta-a|^2$ で推定する統計的決定問題は、 S, F の系列を z として

表9.3 統計的決定理論の一般的な前提

状態	行動	$a_1 \quad a_2 \quad \dots \quad a_K$	$p_{11} \quad p_{12} \quad \dots \quad p_{1L}$	計
		θ_1	θ_2	
θ_1	$l_{11} \quad l_{12} \quad \dots \quad l_{1K}$	θ_1	$p_{11} \quad p_{12} \quad \dots \quad p_{1L}$	1
θ_2	$l_{21} \quad l_{22} \quad \dots \quad l_{2K}$	θ_2	$p_{21} \quad p_{22} \quad \dots \quad p_{2L}$	1
\vdots	$\vdots \quad \vdots \quad \vdots$	\vdots	$\vdots \quad \vdots \quad \vdots$	\vdots
θ_I	$l_{I1} \quad l_{I2} \quad \dots \quad l_{IK}$	θ_I	$p_{I1} \quad p_{I2} \quad \dots \quad p_{IL}$	1

左が損失関数、右がデータの確率分布表である。やや抽象的であるが、すぐ前の表9.2を見れば、その意味はすぐわかるだろう。

$$\alpha = \frac{(z \text{ 中の } S \text{ の回数}) + 1}{(\text{試行回数}) + 2}$$

で与えられる(推定量なので、 $\hat{\theta}$ と記してもよい)。ここで、(9.22), (9.27)を用い、一様分布として $\alpha=\beta=1$ とした。

この原理を、わかり易い形で説明しておこう。統計的決定理論では、決定の間接的な対象 θ は、決定者にとって、対面する状態 state と考えられる。損失関数、データの確率分布は、表9.3のようなものとする。

表9.3の左側が、損失関数 $L(\theta, a)$ である。すなわち

$$(9.32) \quad L(\theta_i, a_j) = l_{ij}$$

また、右側は、データ z の(条件付)確率分布の表で

$$(9.33) \quad f(z_j | \theta_i) = p_{ij}$$

である。統計的決定問題は、基本的考え方としては、各 z に対して、最適な行動 $a=a(z)$ を損失 L の期待値の基準

$$(9.34) \quad E(L) = \sum_j f(z_j | \theta_i) \cdot L(\theta_i, a^*(z_j))$$

が最小になるように、(z ごとに)ルールとして決める。行動 a がデータ z によるとしても、どの z が出るかわからないので、 L は定まらない。そこでデータ z の出方で、期待値をとっている。

しかし、この(9.34)の基準自体は θ に依っている。ネルソン氏の問題でも、 θ_1, θ_2 に応じて、 z の確率分布が

$$(0.60, 0.25, 0.15), (0.20, 0.30, 0.50)$$

と異なる。そこで、 θ_i の事前確率でもう一段、期待値をとって、

$$r = \sum_i w(\theta_i) \left\{ \sum_j f(z_j | \theta_i) \cdot L(\theta_i, a^*(z_j)) \right\}$$

を最小化すべきであろう。この二重和を順序変更するとき、 $w(\theta_i)f(z_j|\theta_i)$ が出るが、これはベイズの定理(9.2a)である。

$$w(\theta_i)f(z_j|\theta_i) = w'(\theta_i|z) \cdot f(z_j)$$

と書けるから($f(z)$ はベイズの定理の分母)，結局

$$r = \sum_j f(z_j) \left\{ \sum_i w'(\theta_i|z_j) \cdot L(\theta_i, a(z_j)) \right\}$$

となり、 z ごとに $\{\cdot\}$ の中の形

$$(9.35) \quad \sum_i w'(\theta_i|z_j) \cdot L(\theta_i, a)$$

を最小化する $a=a(z)$ を求めるべきこと、すなわち最小事後期待損失の基準が導かれる。求められた最適な決定方式 $a(z)$ をベイズ決定方式という。

これは、一般に、 w_1, w_2, \dots, w_I ($w_i \geq 0, \sum_i w_i = 1$) に対して、データをまったく含まない

$$(9.36) \quad \sum_i w_i L(\theta_i, a_j)$$

を最小化する $a^* = a_j$ をあらかじめ求めておき、あと w_i に $w'(\theta_i|z_j)$ をあてはめる形にすると、ほとんど自動的に解ける。(9.37)を最小化することをノード・データ問題といふ。

例 9.7 <ネルソン氏の雨の問題：続> 氏の最適行動を求めてみよう。いま、事前確率を $w(\theta_1) = 1/2, w(\theta_2) = 1/2$ として、(9.36)を(9.37)の方法で解くと、ベイズ決定方式は

z_1 に対しては a_1 , z_2 に対しては a_2 , z_3 に対しては a_3 をとる
というものになる^{*)}。事前確率が他の数値なら、これとは異なったベイズ決定方式になることはいうまでもない。

9.5 ベイズ判別

統計的決定問題で表 9.3 が特別な形をもった場合が、いわゆる分類 classification あるいは判別 discrimination の問題である。

分類の問題とは、客体の観測値 z (一般にはベクトル)によって、その客体が属しているはずの K 種類のカテゴリー(あるいは部分母集団) Π_1, \dots, Π_k のい

*) 巻末松原(1992)参照。

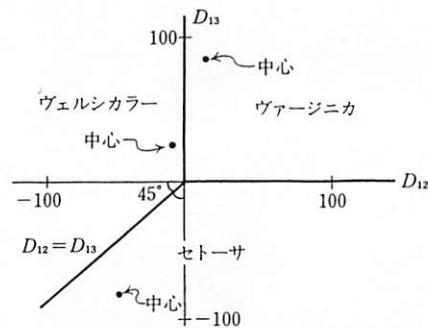
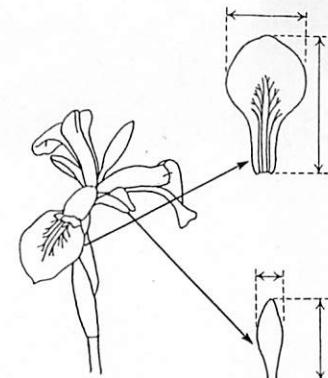


図 9.4 アイリス(アヤメ)の計量分類

D_{12}, D_{13} は線形判別関数。これで判別する。



アイリス(アヤメ)の測定部位

(E. Anderson)

ずれか一つへ、それをできるだけ正しく帰属させることである。判別といつてもほぼ同義である。ただし、分類は最終の行動を、判別は方法・過程を強調する傾向がある。これは、いわゆる線形判別解析をはじめとして、通信工学の信号検出理論、パターン認識理論、また、心理学の閾値理論、人類学、人文地理学など、多くの応用分野の基本的理屈の一つとなっている。

例 9.8 <計量分類> 人間、動物、植物の個体に対する計測(生物計測 biometry)は、F. ゴルトン、K. ピアソン、R. フィッシャー以来、統計学の発祥地とし、統計学の歴史の中でも重要な応用例を与えてきた(第 I 卷第 1 章)。

とりわけ、フィッシャー^{*)}の 3 種のアヤメに対する 4 部位の計測データ(章末)は、アイリス・データ iris data といわれて、多変量解析における判別問題の典型的応用例とされる。扱われた 3 種のアヤメは

Π_1 : ヴァージニカ virginica, Π_2 : ヴェルシカラ versicolor,

Π_3 : セトーサ setosa

であり、4 部位の測定は

x_1 = がく片長さ, x_2 = がく片幅, x_3 = 花びら長さ, x_4 = 花びら幅
である。これは $z = (x_1, x_2, x_3, x_4)$ と 4 次元データになる。

まず、分類の標識を θ とし、 Π_1 では $\theta = \theta_1$, Π_2 では $\theta = \theta_2$ 等々、となつてゐるとする。 $\Theta = \{\theta_1, \dots, \theta_K\}$ とすれば、これが状態の集合である。行動は、

*) 測定データは、Edgar Anderson のものである。

表9.4 カテゴリーの分類の損失関数

θ	a	a_1	a_2	\dots	a_k
θ_1		0	1	\dots	1
θ_2		1	0	\dots	\vdots
\vdots		\vdots	1	\ddots	\vdots
\vdots		\vdots	\vdots	\ddots	1
θ_K		1	1	\dots	1
					0

Π_1 に分類(せよ)を a_1 , Π_2 に分類を a_2 等々とする。真の所属先が Π_1 のときのデータの確率分布を $f_1(z)$, 同じく Π_2 のとき $f_2(z)$ 等々とする^{*)}と, あとは損失関数を与えれば, 典型的な統計的決定問題となる。もちろん, z は一次元でも, 多次元でもよい。

いま Π_k に関するデータ z を Π_k へ分類すれば (correct classification) 損失 = 0, $\Pi_l (l \neq k)$ への分類過誤の(misclassification) 損失 = 1 であるものとすると, 損失関数は表 9.4 の形になる(0-1型単純損失)。この統計的応用は次のようにある。確率分布の族 $\{f_\theta\}$ を考え, K 個の母集団 $\Pi_k (k=1, \dots, K)$ では $\theta = \theta_k$ となっているとする。

まず, ノードデータ問題を解く。 $w = (w_1, \dots, w_K)$ を事前分布とすれば a_k に付随する(9.36)の量は

$$(9.37) \quad w_1 + \dots + w_{k-1} + w_{k+1} + \dots + w_k = 1 - w_k$$

となる。よってすべての $l (l=1, \dots, K)$ に対して

$$(9.38) \quad 1 - w_k \leq 1 - w_l$$

ならば, a_k をとる(Π_k へ分類する)。これは, $w_k \geq w_l$ である。すなわち, 最大の事前確率をもつ Π に分類する。

よって, z を観測したときは事後確率最大のものへ分類する。ここで Π_k のもつ事後確率は

$$(9.39) \quad w_k' = \frac{w_k f_k(z)}{w_1 f_1(z) + \dots + w_K f_K(z)}$$

である。そこで, z を観測して

$$(9.40) \quad w_k f_k(z) \geq w_1 f_1(z), \dots, w_K f_K(z)$$

ならば Π_k へ分類するのが, ベイズ判別的方式である。尤度比の形にすれば

$$(9.41) \quad \frac{f_1(z)}{f_k(z)} \leq \frac{w_k}{w_1}, \quad \frac{f_2(z)}{f_k(z)} \leq \frac{w_k}{w_2}, \dots, \quad \frac{f_K(z)}{f_k(z)} \leq \frac{w_k}{w_K}$$

ならば, データ z を第 k 部分母集団 Π_k へ分類するという形になる。これが分類の条件式である。

*) 印刷の都合上, $f(z|\theta_i)$ を $f_i(z)$ と記す。

表9.5[†] フィッシャーによるアイリス・データ: 3種のアヤメの4部位の測定値(図9.4参照)

アイリス・ヴァージニカ				アイリス・ヴェルシカラ				アイリス・セトーサ			
がく片長さ	がく片幅	花弁長さ	花弁幅	がく片長さ	がく片幅	花弁長さ	花弁幅	がく片長さ	がく片幅	花弁長さ	花弁幅
6.3	3.3	6.0	2.5	7.0	3.2	4.7	1.4	5.1	3.5	1.4	0.2
5.8	2.7	5.1	1.9	6.4	3.2	4.5	1.5	4.9	3.0	1.4	0.2
7.1	3.0	5.9	2.1	6.9	3.1	4.9	1.5	4.7	3.2	1.3	0.2
6.3	2.9	5.6	1.8	5.5	2.3	4.0	1.3	4.6	3.1	1.5	0.2
6.5	3.0	5.8	2.2	6.5	2.8	4.6	1.5	5.0	3.6	1.4	0.2
7.6	3.0	6.6	2.1	5.7	2.8	4.5	1.3	5.4	3.9	1.7	0.4
4.9	2.5	4.5	1.7	6.3	3.3	4.7	1.6	4.6	3.4	1.4	0.3
7.3	2.9	6.3	1.8	4.9	2.4	3.3	1.0	5.0	3.4	1.5	0.2
6.7	2.5	5.8	1.8	6.6	2.9	4.6	1.3	4.4	2.9	1.4	0.2
7.2	3.6	6.1	2.5	5.2	2.7	3.9	1.4	4.9	3.1	1.5	0.1
6.5	3.2	5.1	2.0	5.0	2.0	3.5	1.0	5.4	3.7	1.5	0.2
6.4	2.7	5.3	1.9	5.9	3.0	4.2	1.5	4.8	3.4	1.6	0.2
6.8	3.0	5.5	2.1	6.0	2.2	4.0	1.0	4.8	3.0	1.4	0.1
5.7	2.5	5.0	2.0	6.1	2.9	4.7	1.4	4.3	3.0	1.1	0.1
5.8	2.8	5.1	2.4	5.6	2.9	3.6	1.3	5.8	4.0	1.2	0.2
6.4	3.2	5.3	2.3	6.7	3.1	4.4	1.4	5.7	4.4	1.5	0.4
6.5	3.0	5.5	1.8	5.6	3.0	4.5	1.5	5.4	3.9	1.3	0.4
7.7	3.8	6.7	2.2	5.8	2.7	4.1	1.0	5.1	3.5	1.4	0.3
7.7	2.6	6.9	2.3	6.2	2.2	4.5	1.5	5.7	3.8	1.7	0.3
6.0	2.2	5.0	1.5	5.6	2.5	3.9	1.1	5.1	3.8	1.5	0.3
6.9	3.2	5.7	2.3	5.9	3.2	4.8	1.8	5.4	3.4	1.7	0.2
5.6	2.8	4.9	2.0	6.1	2.8	4.0	1.3	5.1	3.7	1.5	0.4
7.7	2.8	6.7	2.0	6.3	2.5	4.9	1.5	4.6	3.6	1.0	0.2
6.3	2.7	4.9	1.8	6.1	2.8	4.7	1.2	5.1	3.3	1.7	0.5
6.7	3.3	5.7	2.1	6.4	2.9	4.3	1.3	4.8	3.4	1.9	0.2
7.2	3.2	6.0	1.8	6.6	3.0	4.4	1.4	5.0	3.0	1.6	0.2
6.2	2.8	4.8	1.8	6.8	2.8	4.8	1.4	5.0	3.4	1.6	0.4
6.1	3.0	4.9	1.8	6.7	3.0	5.0	1.7	5.2	3.5	1.5	0.2
6.4	2.8	5.6	2.1	6.0	2.9	4.5	1.5	5.2	3.4	1.4	0.2
7.2	3.0	5.8	1.6	5.7	2.6	3.5	1.0	4.7	3.2	3.6	0.2
7.4	2.8	6.1	1.9	5.5	2.4	3.8	1.1	4.8	3.1	1.6	0.2
7.9	3.8	6.4	2.0	5.5	2.4	3.7	1.0	5.4	3.4	1.5	0.4
6.4	2.8	5.6	2.2	5.8	2.7	3.9	1.2	5.2	4.1	1.5	0.1
6.3	2.8	5.1	1.5	6.0	2.7	5.1	1.6	5.5	4.2	1.4	0.2
6.1	2.6	5.6	1.4	5.4	3.0	4.5	1.5	4.9	3.1	1.5	0.2
7.7	3.0	6.1	2.3	6.0	3.4	4.5	1.6	5.0	3.2	1.2	0.2
6.3	3.4	5.6	2.4	6.7	3.1	4.7	1.5	5.5	3.5	1.3	0.2
6.4	3.1	5.5	1.8	6.3	2.3	4.4	1.3	4.9	3.6	1.4	0.1
6.0	3.0	4.8	1.8	5.6	3.0	4.1	1.3	4.4	3.0	1.3	0.2

表9.5 つづき

アイリス・セトーサ				アイリス・ヴェルシカラー				アイリス・ヴァージニカ			
がく片 長さ	がく片 幅	花 弁 長さ	花 弁 幅	がく片 長さ	がく片 幅	花 弁 長さ	花 弁 幅	がく片 長さ	がく片 幅	花 弁 長さ	花 弁 幅
6.9	3.1	5.4	2.1	5.5	2.5	4.0	1.3	5.1	3.4	1.5	0.2
6.7	3.1	5.6	2.4	5.5	2.6	4.4	1.2	5.0	3.5	1.3	0.3
6.9	3.1	5.1	2.3	6.1	3.0	4.6	1.4	4.5	2.3	1.3	0.3
5.8	2.7	5.1	1.9	5.8	2.6	4.0	1.2	4.4	3.2	1.3	0.2
6.8	3.2	5.9	2.3	5.0	2.3	3.3	1.0	5.0	3.5	1.6	0.6
6.7	3.3	5.7	2.5	5.6	2.7	4.2	1.3	5.1	3.8	1.9	0.4
6.7	3.0	5.2	2.3	5.7	3.0	4.2	1.2	4.8	3.0	1.4	0.3
6.3	2.5	5.0	1.9	5.7	2.9	4.2	1.3	5.1	3.8	1.6	0.2
6.5	3.0	5.2	2.0	6.2	2.9	4.3	1.3	4.6	3.2	1.4	0.2
6.2	3.4	5.4	2.3	5.1	2.5	3.0	1.1	5.3	3.7	1.5	0.2
5.9	3.0	5.1	1.8	5.7	2.8	4.1	1.3	5.0	3.3	1.4	0.2

アヤメ科は寒帯から熱帯まで広く分布し、約70属1800種を含み、特にアヤメ属(アイリス、ハナショウブなど)、サフラン属(クロッカス)、フリージア属、グラジオラス属などが日本では普及している。そのうち、アヤメ属は北半球温帯に200種が広く分布、多くが鑑賞用に栽培される。日本ではカキツバタがアヤメ属としては最も古くから知られる(平凡社『世界百科事典』)。

本データは3種類各50を4部位について計測した結果である。フィッシャーによって、計量分類学の例として紹介されて以来、判別分析など広く多変量解析の分析例として用いられてきたもの。

例 9.9 (9.41) の f_1, f_2, f_3 には、4次元正規分布と仮定し(第I巻第7章、本書第2章参照)、 x_1, x_2, x_3, x_4 が確率分布において平均だけ異なるとする。多次元正規分布は、 x_1, x_2, x_3, x_4 の2次式の指數関数(exp)であるから平均および分散共分散行列を推定して、分類の条件式(9.42)へ代入すると、次の式を得る。すなわち

$$(9.42) \quad \begin{aligned} D_{12}(z) &= -3.2456x_1 - 3.3907x_2 + 7.5530x_3 + 14.6358x_4 - 31.5226 \\ D_{13}(z) &= -11.0759x_1 - 19.9161x_2 + 29.1874x_3 + 38.4608x_4 - 18.0933 \end{aligned}$$

として、 z に対して

- i) $D_{12} > 0, D_{13} > 0$ ならば、 Π_1 (ヴァージニカ)へ、
- ii) $D_{12} < 0, D_{23} \equiv D_{13} - D_{12} > 0$ ならば、 Π_2 (ヴェルシカラー)へ、
- iii) $D_{13} < 0, D_{23} < 0$ ならば、 Π_3 (セトーサ)へそれぞれ分類する。

これらの $D(z)$ を線形判別関数 linear discriminant function という。これらの判別の条件 1)~iii) は、コンピューターにまかせることができるので、植物分類のみならず、生物学一般、医学、心理学、人文地理学、経営学などに広く

用いられている。この方法を、計量分類 numerical taxonomy という。

ランドサットなどの衛星写真を、色、明度に転換し、地上の土地利用、植生を判別する技術は、リモート・センシングの一方法であるが、これも計量分類の応用の分野である。

◆◆◆◆◆ 第9章 練習問題 ◆◆◆◆◆

9.1 (ベイズの定理による確率の更新) 3つのつぼ H_1, H_2, H_3 があり、それぞれ、 $3:1, 1:1, 2:1$ の比率で、 S, F と記された玉が入っている。ランダムに指定されたつぼから、玉を復元抽出で抜いたとき、玉は S であった。

- i) その玉がつぼ $H_i (i=1, 2, 3)$ からのものである事後確率 $P(H_i|S)$ は、それぞれいくつになるか。事前確率は $P(H_1)=P(H_2)=P(H_3)=1/3$ とする。
- ii) 指定された同じつぼから、もう1回玉を抜いたとき、再び S であった。回を一般に添字 $_{1,2}$ で表して、これを S_2 と表そう。この時点でのつぼ H_i の事後確率 $P(H_i|S_1 S_2)$ を求めよ。
- iii) 各 H_i に対し、 $S_1 S_2$ の確率を考える方法で、ii) を求めよ。

9.2 (予測分布) 9.1のi) と同一の状況において、次に S (つまり S_2) となる確率 $P(S)$ を、 $P(S|H_i) (i=1, 2, 3)$ の期待値として求めよ。 $P(S), P(F)$ を観測値の予測分布 predictive distribution という。

9.3 (同、ベータ分布の場合) 成功 S 、失敗 F の二項確率を $P(S)=\theta, P(F)=1-\theta (0 \leq \theta \leq 1)$ としよう。当初、 θ の事前確率分布は一様分布とする。いま、5回までの逐次の試行の結果が S, F, S, S, S であったとするとき、6回目の成功、失敗の結果 S, F の確率分布を、5回目における事後確率分布 $Be(5, 2)$ の密度関数 $w'(\theta|z)$ から

$$\int_0^1 \theta \cdot w'(\theta|z) d\theta, \quad z=(4, 1)$$

によって計算せよ。

9.4 (正規分布の自然な共役分布) ある量 X の母集団分布が $N(\theta, 1/4)$ のとき、 θ を推定するために θ の事前確率分布として平均 5、分散 1/2 の正規分布 $N(5, 1/2)$ を仮定した。 X の観測値が $X=6.5$ のとき、 θ の事後確率分布を求めよ。

9.5 (ベイズ診断) ある疾病 D (たとえば、ぜんそく)があるかどうかを診断するために、3つの症候 S_1, S_2, S_3 があるかどうかを調べるものとする。 S_1, S_2, S_3 はそれぞれ確率的に生じ、かつ独立とし、また、 D の場合と D でない場合 \bar{D} (正常など

き)とでは、これらの症候の確率(尤度)が異なり、以下の表で与えられている。なお、疾病 D の発生確率(統計的な発生率)も事前に与えられている。

状態	発生(確)率	各症候の確率(尤度)		
		S_1	S_2	S_3
$D(\text{ぜんそく})$	0.23	0.1	0.7	0.6
$\bar{D}(\text{ぜんそくでない})$	0.77	0.8	0.2	0.5

- i) 症候 S_3 はなく、 S_1, S_2 があるとき、このパターンを $S=(1, 1, 0)$ で表そう。このときの、 D の事後確率 $P(D|S)$, \bar{D} の事後確率 $P(\bar{D}|S)$ を求めよ。
- ii) 比 $\Omega(S)=P(D|S)/P(\bar{D}|S)$ を、8通りのすべてのパターン S について計算し、大きさの順に並べよ。また、この比を最大にする(もっとも高い可能性で‘ぜんそく’と診断される)症候のパターンはどれか。比 $\Omega(S)$ を、 D の事後オッズ posterior odds という。
- iii) 事後オッズ $\Omega(S)$ の大きさの順序は、尤度比 $\lambda(S)=P(S|D)/P(S|\bar{D})$ の順序に一致することを示せ。また、 $\lambda(S)\geq 1$ となるパターン S をすべて挙げよ。

9.6 <統計的決定理論> ネルソン氏の状況(表 9.2)において、

- i) 晴天 θ_1 , 雨天 θ_2 の事前確率が $1/2, 1/2$ のとき、ネルソン氏にとって a_1, a_2, a_3 のどの行動が最適か。一般に、事前確率が $w, 1-w$ のときはどうか。 w の範囲で分類して答えよ。また、晴天用で外出するべきときの w の範囲を求めよ。(ノード・データ問題)
- ii) 晴雨計の示度のデータがあるとして、示度が $z_i (i=1, 2, 3)$ のとき、表 9.2 右表から晴天 θ_1 , 雨天 θ_2 の事後確率を各 z_i に対してを求めよ。
- iii) i)の事前確率 $w, 1-w$ を ii)で求めた事後確率に代えて、 a_1, a_2, a_3 の中からネルソン氏の最適行動を決めよ。

9.7 <線形判別関数による計量分類> アイリス・データの各種 10 ケース計 30 ケースをランダムに選び出し、それを(9.37)の線形判別関数で分類してみよ。分類の過誤(Π_i に属するケースが $\Pi_j (j \neq i)$ へ分類されること)は生じているか。

9.8 <リンドレーの 3 囚人問題> 次のベイズの定理の適用につきコメントせよ。
——3人の囚人が幽閉されているとしよう。3人の囚人の名前は、アラン(Alan), バーナード(Bernard), チャールズ(Charles)とする。アランは、翌日3人のうち2人が処刑され、1人が釈放されることを知っているが、3人のうち誰が釈放されるかについては全くわからない。このような状況において、アランが看守に対し、「3人のうち2人が処刑されるのは確実である。バーナードとチャールズのうち少なくとも1人は処刑されるのは確実である。バーナードとチャールズのうち処刑される者の名前を1人だけ教えてくれても、アランの釈放については全く情報を与えないは

ずだから、その名前を教えてほしい。」と言ったところ、看守は、アランの言い分に納得し、「バーナードは処刑される。」と答えた。アランはこれを聞いて、釈放される可能性があるのは、自分その他はチャールズのみになったので、自分が釈放される可能性が増えたと喜んだとする。直観的には正しいように見えるが、この確率評価は根拠があるだろうか。

釈放されるべき人を、頭文字で、 A, B, C 、とし、また s を「バーナードは処刑される」という言明とすると、状況から、 $P(s|A)=1/2, P(s|B)=0, P(s|C)=1$ と仮定される。事前確率は、 $P(A)=P(B)=P(C)=1/3$ とする。ベイズの定理から

$$\begin{aligned} P(A|s) &= \frac{P(A)P(s|A)}{P(A)P(s|A)+P(B)P(s|B)+P(C)P(s|C)} \\ &= \frac{(1/3) \cdot (1/2)}{(1/3) \cdot (1/2) + (1/3) \cdot 0 + (1/3) \cdot 1} \\ &= 1/3 \end{aligned}$$

ゆえに、 $P(A)=(A|s)$ となり、上記の確率評価は根拠がない。——

9.9 <信頼性のベイズ的評価> i) 指数分布、ガンマ分布の再生性 $Ex(\lambda)*Ex(\lambda)=Ga(2, \lambda)$ を証明せよ。

ii) ある工業用エンジンの故障時間は平均 θ の指数分布 $Ex(1/\theta)$

$$f(x)=(1/\theta)e^{-x/\theta}$$

に従うとしよう。2人の設計者 A 氏、B 氏は共同でエンジンの改良設計を行うことになった。仕様によれば、 $\theta \geq 3000(h)$ とならねばならない。ところで、a)過去に類似の実績のある A 氏によれば、このことは確率 0.75 で可能であるが、 $\theta \leq 2000$ となる失敗ケースも小さい確率 0.05 でありうると思われた。また、b)B 氏は A 氏ほど経験がなく、このような失敗は確率 0.30 で起ると考えた。

A 氏は比較的楽観的、B 氏は悲観的見方をしているわけであるが、くわしくいえば、A 氏、B 氏の、 θ の事前確率分布 $w_A(\theta), w_B(\theta)$ は下表のようである。

表 設計者 A, B の、平均 θ の事前確率分布
 $w_A(\theta), w_B(\theta)$

平均 θ の範囲 (区間内では一様)	事前確率	
	設計者 A	設計者 B
0—1000(h)	0.01	0.15
1000—2000	0.04	0.15
2000—3000	0.20	0.20
3000—4000	0.50	0.20
4000—5000	0.15	0.15
>5000	0.10	0.15
	1.00	1.00

$w_A(\theta)$, $w_B(\theta)$ のグラフを, θ (横軸)の単位を $1000 \text{ h.} = 1$ として^{*)}それぞれ書け. ただし, >5000 の階級は, 便宜的に $5000\text{-}6000$ と考えよ.

iii) 改良した試作型エンジン 2 機をテストして, 故障時間が, $x_1=2000$, $x_2=2500$ であった. この標本和 $t=x_1+x_2=4500$ であるとき, A, B の, θ の事後確率分布

$$w'_A(\theta|2000, 2500), \quad w'_B(\theta|2000, 2500)$$

を表と同様の形に求めよ.

なお, 微積分学の知識から,

$$\int_a^b \frac{t}{\theta^2} e^{-t/\theta} d\theta = e^{-t/b} - e^{-t/a} \quad (t > 0, b \geq a \geq 0)$$

を用いてよい. $a=0$ のときは, 右辺第 2 項 = 0 とする.

第 10 章 確率過程の基礎

世の中の現象は, よく見てみると, 自然現象でも社会現象でもひとつ所にとどまらず, 時の経過につれて変化する. 先を予測することは易しいことではない. そこで, 不確実な将来的予測・制御をするために, 「変化」を確率過程としてモデル化してみよう.

数々の興味深い応用例もある.

10.1 ランダム・ウォークと破産問題

10.1.1 確率過程

天候, 地震の発生, 生物学における遺伝, 経済であれば株価, 為替レートなど, 時間の推移とともに刻々と変化していくさまざまな現象がある. このような現象の変化の過程は, ランダムな確率法則に支配されていると見なして分析すると有効な場合が多い. いま各時点で観測されるデータは確率変数の実現値と見なす. 午前 0 時, 1 時の気温などと, 離散的に観測する場合は, 時点を n ($n=0, \pm 1, \pm 2, \dots$), 観測値を X_n , また地震波, 脳波などのように, 連続的に観測する場合は, 時点を t ($-\infty < t < \infty$), 観測値を X_t と記す. ただし n, t はそれぞれ整数, 実数の一部分を動くこともある. このとき確率変数の集合 $\{X_n\}, \{X_t\}$ を確率過程 stochastic process という. また, X が実現値のときは, n (または t) の関数となる. このとき $\{X_n\}, \{X_t\}$ を標本関数 sample function, サンプル・パス sample path ということがある(なお, 観測方法(時間)の離散・連続と, 確率分布の離散・連続とは無関係である. たとえば連続的に観測しても, 時点ごとの確率分布は離散的な場合もある).

*) 作図上のものでなく, w が確率分布になるためである. 結果的には単位は任意である.

これまで互いに独立で同一な分布に従う確率変数列を考察してきた。これらも確率過程の一例であるが、ここではさらに発展させ、異時点の確率変数の間に相関関係があったり、分布が時点ごとに変化していく確率過程も扱う。

目的や対象となる現象の複雑さの度合に応じて、いろいろな確率過程が提案されているが、ここでは基本的なものとして、ランダム・ウォーク、ブラウン運動、マルコフ連鎖、その代表例としてポアソン過程、出生死滅過程について説明する。応用例として、さまざまなサービスを受けるときの待ち時間、待ち行列の確率的な表現、株価データの解析を紹介する。

10.1.2 ランダム・ウォーク

一つの粒子がある直線に沿って、上下に(数直線では左右に)運動を繰り返しているとしよう。時点、 $0, 1, 2, \dots$ で観測し、簡単のため各時点で確率 p で上へ 1 単位、確率 $q = 1 - p$ で下へ 1 単位移動すると仮定する。時点 0 における位置を 0 とし、時間の推移とともに変化する粒子の位置を表したのが図 10.1 である。これはランダム・ウォーク random walk と呼ばれている。まさに、「酔っぱらい」の千鳥足に似ているので酔歩とよばれることもある。ただし注意しなくてはならないのは、図 10.1 の時点 n におけるたて軸方向の上下の位置を表しているだけで、粒子は前に進んでいるわけではない。したがって厳密

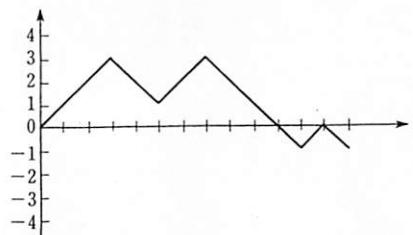


図 10.1' ランダム・ウォークの例
時間の経過に従って粒子は上下いでたらめな運動を繰り返している。これを定式化するのがランダム・ウォークである。

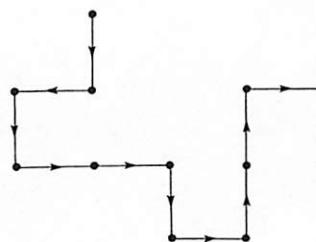


図 10.2 2 次元ランダム・ウォーク
実際に我々が酔っぱらったときの千鳥足は、東西南北にふらつく 2 次元のランダム・ウォークである。

には、1次元のランダム・ウォーク 1-dimensional random walk とよぶべきものである。真の酔っぱらいの挙動を示すには、図 10.2 のような上下左右にふらつく 2 次元のランダム・ウォーク 2-dimensional random walk を考える必要がある。さらには任意の次元のランダム・ウォークを考えることも可能である。

以下では1次元のランダム・ウォークに限定する。いま X_1, X_2, \dots, X_n は互いに独立で、それぞれ確率 p で 1 をとり、確率 q で -1 をとる確率変数とする。時点 n における粒子の位置は、

$$S_n = X_1 + X_2 + \cdots + X_n$$

と表現できる。ここで時点 n において、位置 r にいる確率を求めよう。いま $(X_i+1)/2$ は、ベルヌーイ試行列になる。したがって $B_n = (S_n + n)/2$ とすれば、 B_n の確率分布は二項分布である。第 1 章の結果を用いて

$$P(S_n=r) = P(B_n=(n+r)/2) \\ = {}_n C_{(n+r)/2} p^{(n+r)/2} q^{(n-r)/2}$$

$(n+r)$ が偶数, かつ $0 \leq (n+r)/2 \leq n$ のとき

を得る。次に原点を出発したランダム・ウォークが、原点に戻ってくる確率は、
 $P(\text{少なくとも一つの } n=1, 2, \dots \text{ に対して, } S_n = 0)$

$$1 - |p - q|$$

に等しいことが知られている。つまり $p=q=1/2$ のときは、「酔っぱらい」はいつかは必ず原点に戻ってくる。しかしそし $p \neq q$ のときは、二度と戻ってこないこともある。このことは、つぎの事実から想像できる。 X_i の期待値は $(p-q)$ であるから、大数の強法則より S_n/n は、確率 1 で $p-q$ に収束する。したがって S_n 自身は、最終的に $p > q$ のときは正の無限大 ($+\infty$) に、 $p < q$ のときは負の無限大 ($-\infty$) に発散してしまうからである。

10.1.3 破産問題

ランダム・ウォークに関連した例として、ここでは古くから有名な破産問題 ruin problem を取り上げよう。

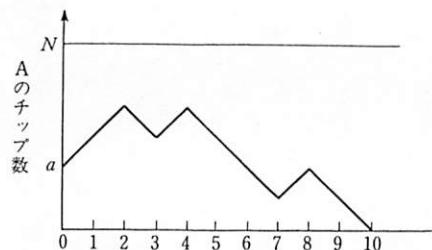


図 10.3 破産問題

ランダム・ウォークの例として有名な破産問題である。コイン・トスの裏表によりチップをやりとりしていく。「裏目」が出すぎて、Aは10回目に破産した。

AとBの2人がいて、最初にAが a 枚、Bが $N-a$ 枚のチップ(一定金額を表す札)を持っていたとしよう。コインを投げ続け、表が出たときにはAがBから1枚チップをもらい、逆に裏が出たときには、AがBに1枚チップをあげるとする。どちらか一方の手持ちのチップがなくなり、「破産」したときゲームは終了とする。図10.3は、序盤戦において、Aは優勢であったが、後半負け続け、最終的には10回目で破産した様子を示している。もし N を到達すれば、Bが破産したことになる。0と N は、吸収壁 absorbing barrierともよばれる。コインの表が出る確率を p 、裏が出る確率を $q (=1-p)$ として、Aの破産する確率を求めよう。

この確率は当然 a によるので、 $r(a)$ でこの破産確率を表そう。いま事象 R 、 H をそれぞれ

$R=\{A\text{ がいつか破産する}\}, \quad H=\{\text{第1回目にコインの表が出る}\}$
と定義する。条件付き確率の性質から

$$P(R)=P(R|H)P(H)+P(R|H^c)P(H^c)$$

である。第1回目にコインの表が出れば、Aのチップの所有枚数は $a+1$ となり、裏が出れば $a-1$ となるので、 $P(R|H)=r(a+1)$ 、 $P(R|H^c)=r(a-1)$ である。したがって $a=1, 2, 3, \dots, N-1$ に対して、差分方程式

$$(10.1) \quad r(a)=pr(a+1)+qr(a-1)$$

が成立する。ここで境界で成り立つべき性質

$$r(0)=1, \quad r(N)=0$$

を条件にして、差分方程式(10.1)を解くと、破産確率 $r(a)$ は

$$(10.2) \quad r(a)=\frac{(q/p)^a-(q/p)^N}{(q/p)^N-1} \quad (p \neq q \text{ のとき})$$

$$(10.3) \quad =1-a/N \quad (p=q \text{ のとき})$$

となる(差分方程式の解法については章末を参照のこと)。

逆にBの破産する確率を $w(a)$ とすれば、これは(10.2)、(10.3)において p と q 、 a と $N-a$ をそれぞれ入れ換えるべきである。

$$\begin{aligned} w(a) &= \frac{(p/q)^N - (p/q)^{N-a}}{(p/q)^N - 1} && (p \neq q \text{ のとき}) \\ &= a/N && (p=q \text{ のとき}) \end{aligned}$$

となる。したがって

$$r(a)+w(a)=1$$

が成立し、この賭けは $p \neq q$ のときはもちろんのこと、 $p=q$ でもかならず勝負がつく。 $p=q$ のとき賭けは公平 fair という。賭けが公平だと勝ちも負けも生じないと想像する人もいるかも知れないが、やはり必ず勝負がつくのである。

このとき A の最終的な利益あるいは損失を G とすれば、 G の期待値つまり期待利益は

$$\begin{aligned} E(G) &= (1-r(a)) \cdot (N-a) + r(a) \cdot (-a) \\ &= N(1-r(a))-a \end{aligned}$$

である。したがって $p=q$ のときは(10.3)より $E(G)=0$ であり、公平な賭けは最初の所持金にかかわらずやはり公平な賭けであることがわかる。

つぎに A あるいは B のどちらかが破産して賭けが終了するまでに、コインを投げた回数を F として、 F の期待値つまり賭けの期待終了時間を探求しよう。これももちろん a に関係する。 $P(R)$ と同様に

$$E(F)=E(F|H)P(H)+E(F|H^c)P(H^c)$$

が成立する。 $e(a)=E(F)$ とおけば

$$E(F|H)=1+e(a+1), \quad E(F|H^c)=1+e(a-1)$$

であるから、 $a=1, 2, \dots, N-1$ に対して、差分方程式

$$(10.4) \quad e(a)=\{1+e(a+1)\}p+\{1+e(a-1)\}q$$

が成立する。ここで

$$e(0)=e(N)=0$$

を境界条件にして、差分方程式(10.4)を解くと、期待終了時間として

$$e(a)=\left(N\frac{(q/p)^a-1}{(q/p)^N-1}-a\right)/(p-q) \quad (p \neq q \text{ のとき})$$

$$=a(N-a) \quad (p=q \text{ のとき})$$

が導ける。

例 10.1 〈破産問題の計算〉 p, N, a に具体的な数値を与えた場合の破産確率 $r(a)$, 期待利得 $E(G)$, 期待終了時間 $e(a)$ は, 表 10.1 の通りである。まず A, B の最初の所持金の比が 9:1 の場合を考えよう。公平な賭け $p=0.5$ の場合には, 破産確率 $r(a)$ は所持金の比 a/N にのみ依存し, A の破産する確率は 0.1 にしかすぎない。賭けの期待終了時間 $e(a)$ は, 比は同じでも合計額 N の 2 乗に比例して大きくなる。

$p=0.45$ と A にとって不利な賭けでも, $r(a)=0.21$ である。ただ 1 回ごとの賭金を小さくすると, これは N を大きくするのと原理は同じであるが, $r(a)=0.87$ と大きくなる。また期待終了時間は $e(a)=766$ と飛躍的に長くなる。

$p=0.4$ とさらに不利な賭けの場合, 所持金の比を 99:1 まで引き上げれば, 破産確率は $r(a)=0.33$ に押さえることができる。ただし $E(G)=-32.3$ となることを覚悟しなくてはならない。

最後に $N=\infty$, つまり相手が無限にお金を持っている場合を考えてみよう。このとき A の破産確率は

$$r(a)=(q/p)^a \quad (p>q \text{ のとき})$$

$$r(a)=1 \quad (p\leq q \text{ のとき})$$

となる。厳密には証明が必要であるが, $\lim_{N\rightarrow\infty} r(a)$ を考えれば想像がつく。

これはラスベガスやモンテカルロのカジノに行くときの参考になろう。不正のな

表 10.1[†] 破産問題の数値例

+1の確率 p	所持金 N	破産確率 $r(a)$	期待利得 $E(G)$	期待終了時間 $e(a)$
0.5	9	0.1	0	9
	90	0.1	0	900
	900	0.1	0	90,000
0.45	9	0.21	-1.1	11
	90	0.87	-76.6	766
0.4	99	0.33	-32.3	162

$p=q=0.5$ であれば $r(a)$ は a/N のみに依存し, $e(a)$ は N^2 に比例する。
 $P<0.5$ のとき $r(a), e(a)$ は N とともに増加する。

い公平な賭けでも, 事実上 $N=\infty$ であるカジノで‘大儲け’しようなどとは, 考えないうが良さそうである。

10.2 ブラウン運動

ランダム・ウォークでは, 観測時点も, 移動する幅も離散的な場合を考えたが, 実際の粒子は, 連続的な時間の流れの中で, 運動も連続的にくり返している。したがって, 時点の間隔と, 位置が変化する幅を同時に細かくしていったときに, ランダム・ウォークが極限においてどのように表現できるか考えよう。

まずランダム・ウォーク S_n の期待値と分散は

$$(10.5) \quad E(S_n)=nE(X_i)=n\{1\cdot p+(-1)\cdot q\}=n(p-q)$$

$$(10.6) \quad V(S_n)=nV(X_i)=n\{p+q-(p-q)^2\}=4npq$$

である。ここで 1 回に変化する幅を Δr 単位, それに要する時間を Δt と細かくしよう。このとき時間 t の間に粒子は $t/\Delta t$ 回動く。したがって(10.5), (10.6)において n を $t/\Delta t$ に, 幅 1 単位を Δr 単位に置き換えれば, 時間 t までの粒子の総変位の期待値と分散は, それぞれ

$$(p-q)t(\Delta r/\Delta t), \quad 4pqt(\Delta r)^2/\Delta t$$

となる。さて Δr と Δt を 0 に近づけたいのであるが, 両方を勝手に動かすことはできない。なぜならまず分散に注目すると, これが一定値に収束するには, $(\Delta r)^2/\Delta t$ が極限値を持たなくてはならない。ただしこのままだと, $\Delta r/\Delta t$ は発散しなければならず, 期待値の極限が存在しない。そこで $p-q$ を Δr と同じ速度で 0 へ収束させて, 期待値の極限も存在するようにする。以上を

$$2D=\lim_{\Delta t, \Delta r \rightarrow 0} \frac{(\Delta r)^2}{\Delta t}, \quad p=1/2+\frac{C\Delta r}{2D}, \quad q=1/2-\frac{C\Delta r}{2D}$$

と表す。 D と C は定数で, それぞれ D は拡散係数 diffusion coefficient, C はすれ drift とよばれている。このとき期待値と分散は, それぞれ

$$(p-q)t \cdot \frac{\Delta r}{\Delta t} = t \left(\frac{C}{D} \right) \cdot \frac{(\Delta r)^2}{\Delta t}$$

$$4pqt \cdot \frac{(\Delta r)^2}{\Delta t} = 4 \left(\frac{1}{2} - \frac{C\Delta r}{2D} \right) \left(\frac{1}{2} + \frac{C\Delta r}{2D} \right) t \cdot \frac{(\Delta r)^2}{\Delta t}$$

$$= t \cdot \frac{(\Delta r)^2}{\Delta t} - \left(\frac{C \Delta r}{D} \right)^2 t \cdot \frac{(\Delta r)^2}{\Delta t}$$

となる。

ここで Δr と Δt を 0 に近づけると、時点 t までの総変位の期待値と、分散の極限値は、それぞれ $2Ct$ と $2Dt$ になる。また、独立で同一分布に従う確率変数の無限和になっているので、中心極限定理により正規分布に従う。したが

※賭博者が賭けから学ばなかった最も大切なこと※

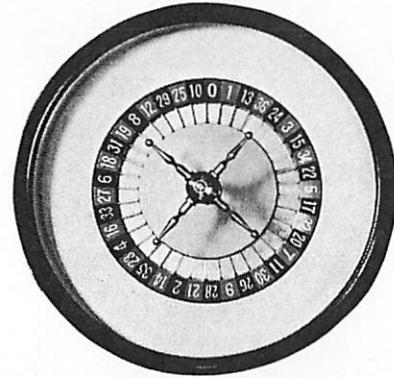
ウィーバー (W. Weaver 1894-) は、シャノン (C. Shannon 1916-) とともに、世紀の名著『情報の数学理論』Mathematical Theory of Communication (1948) を著して、一躍、情報理論の創始者として有名となった人であるが、こう言っている：

「確証を与えることはできない
が、信頼すべき情報によると、モ
ンテ・カルロで昔“偶数”が 28 回
つづけて出たということである。
モンテ・カルロでは、ルーレット
の輪には 1 から 36 までと 0 の数
字がはめこまれていて、数字 0 は
奇数でも偶数でもないと定めら
れている。こうして球が偶数の上
に止まるチャンスは $\frac{18}{37}$ 、つまり
半分よりわずか少ない。換言する

と、つづけて 28 回“偶数”であるのは 1 枚の貨幣を投げたとき 28 回つづけて表が出るのとほとんど同じである。この確率は 1 を 268435456 で割ればよい。
モンテ・カルロでは各テーブルで毎回“偶数”について 3 種類の独立なプレ
イを、1 テーブルにつき 500 回行うから、平均して、1 テーブルにつき、ほぼ
500 年に 1 度、28 回つづけて“偶数”が出なければならない。」

といっても、そんなに賭けをやってもいいのだろうか。

「賭博者は疑いもなく確率論から多くの事柄を学んだが、確率論が賭博につ
いて到達し得る主たる教訓は決して学ばなかった——あるいはいずれにして
も、受け入れたことがない——ことは明白である。その教訓は簡単にいうと、
『賭けをし続けると損をする』ということだからである。」



って時点 t における粒子の位置は、期待値 $2Ct$ 、分散 $2Dt$ の正規分布 $N(2Ct, 2Dt)$ に従う。これをブラウン運動 Brownian motion とよび B_t と書く。

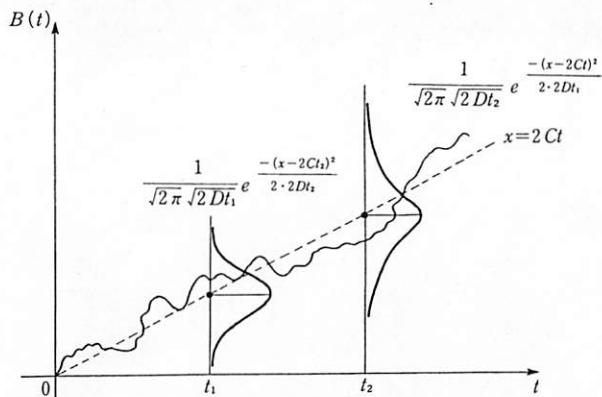


図 10.4 ブラウン運動の挙動

ランダム・ウォークで、観測時点、変化の幅共に連続化したのがブラ
ウン運動である。期待値は直線上にあり、分散は時間と共に増えていく。

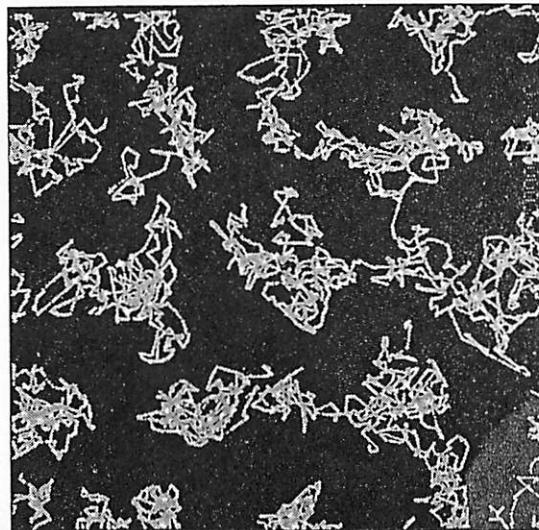


図 10.5 コンピューターがシミュレートした分子運動

二次元のブラウン運動を思わせる。

(出典: Scientific American)

名前の由来は、1826年、植物学者ブラウン(Brown, R. 1773-1856)が、植物の花粉が液体の表面で不規則な運動をする現象として発見したことによる。またのちに1905年、アインシュタイン(Einstein, A.)が定式化したことは有名である。図10.4に示したように、ブラウン運動 B_t は期待値が傾向線 $2Ct$ の上に乗り、分散が比例定数 $2D$ で t に比例して増えていく。このことから D を拡散係数、 C をずれとよぶこともうなづけよう。

ところで時点 s から時点 t ($0 < s < t$) の間の増分 increment $B_t - B_s$ は、同様に期待値 $2C(t-s)$ 、分散 $2D(t-s)$ の正規分布に従う。したがって

$$(10.7) \quad \begin{aligned} \text{Cov}(B_t, B_s) &= (1/2)\{V(B_t) + V(B_s) - V(B_t - B_s)\} \\ &= D\{t+s-(t-s)\} = 2Ds \quad (s < t) \end{aligned}$$

となる。さらに $t_1 < t_2 \leq t_3 < t_4$ として、異時点間の増分 $B_{t_2} - B_{t_1}$ と $B_{t_4} - B_{t_3}$ の共分散を考えよう。(10.7)より

$$\begin{aligned} &\text{Cov}(B_{t_2} - B_{t_1}, B_{t_4} - B_{t_3}) \\ &= \text{Cov}(B_{t_2}, B_{t_4}) - \text{Cov}(B_{t_2}, B_{t_3}) - \text{Cov}(B_{t_1}, B_{t_4}) + \text{Cov}(B_{t_1}, B_{t_3}) \\ &= 2D(t_2 - t_2 - t_1 + t_1) = 0 \end{aligned}$$

となり、無相関である。 $B_{t_2} - B_{t_1}$ と $B_{t_4} - B_{t_3}$ は正規分布に従うので、これは独立性を意味する。したがって、ブラウン運動は独立増分 independent increment の正規過程 normal processともいわれる。

10.3 マルコフ連鎖

10.3.1 推移確率

サイコロ投げ、コイン・トス(コイン投げ)などは、独立な確率変数列の代表例である。だが現実に観測するデータの多くは過去の値の影響を受け、独立な確率変数列とは見なしがたい。そこで独立性の仮定をはずした、確率過程を組み立てることを考えよう。ここでは時点 0 から離散的に観測し、また各時点での確率分布も離散的とする。とりうる値を状態 state とよび、今後ことわりのないかぎり $1, 2, \dots, N$ とする。

独立でない場合、将来の観測値は現在および過去の値の影響を受けるが、遠い過去になるに従いその影響の度合は薄れると仮定するのは自然である。さら

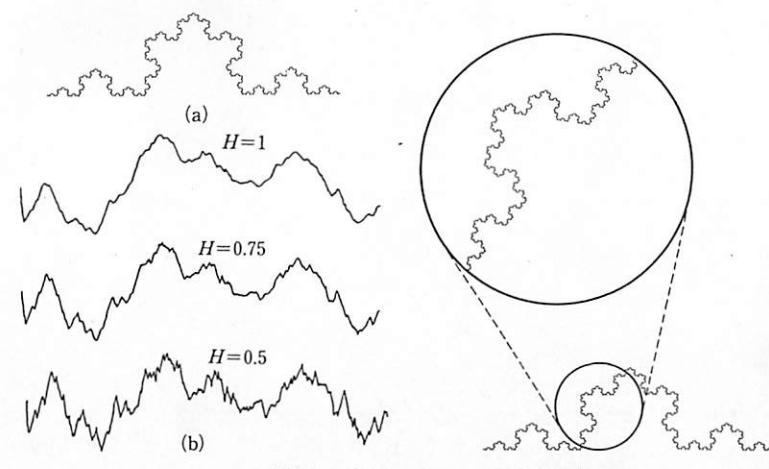
※フラクタル※

図(a)を見よう。コッホ曲線とよばれているがなにやら奇妙である。この図を顕微鏡で10倍、100倍と、限界なく拡大して覗いたとしよう(右図)。再び原型と同じ形状の図が見えるはずである。この性質を自己相似性 self-similarity という。もし微分可能な曲線であれば、倍率を大きくしていくばだんだん滑らかになり直線に見えてくるはずである。一方このように微分不可能で、いくら拡大してもギザギザのある曲線、あるいは空間图形を、マンデルブロー(Mandelbrot, B.)はフラクタル fractal と名づけた。fraction(破片)、fractional(分数の)などからヒントを得た造語である。リアス式の海岸線、峨々たる山の稜線、樹木の葉脈など自然界には多く存在する。

コッホ曲線は通常の関数であるが、分布が自己相似性を持つ確率過程として、フラクショナル・ブラウン運動がある。これは正規過程で

$$V(B_t - B_s) = 2D|t-s|^{2H} \quad (s \leq t)$$

となる。 $H=1/2$ のときが通常のブラウン運動である。 $H \neq 1/2$ のときは、異時点間の増分は独立ではない。図(b)は H をさまざまな値に変化させたときのひとつの実現値の系列である。ここでは $H=0$ としてある。もちろん標本関数自身は、時と場合によってさまざまな挙動を示すので、コッホ曲線のように厳密な自己相似性を持たない。しかしどとなく似たような動きを示している。近年フラクタルは、コンピュータ・グラフィックス、自然科学あるいは経済学におけるデータ解析などで注目を浴びている。



に簡単化して、確率過程 $\{X_n\}$ において、将来の確率変数に対する条件付き確率分布が現在の値のみに依存するとき、つまり任意の n 、任意の状態 $j_0, j_1, \dots, j_{n-1}, j, k$ に対して

$$P(X_{n+1}=k | X_0=j_0, X_1=j_1, \dots, X_{n-1}=j_{n-1}, X_n=j) \equiv P(X_{n+1}=k | X_n=j)$$

が成立するとき、この確率過程はマルコフ性 Markov property をもつという。また $\{X_n\}$ をマルコフ連鎖 Markov chain とよぶ。有限個の過去の値にも依存するときは多重マルコフ連鎖 multiple Markov chain とよぶ。また条件付き分布が n に依存しないときは齊時マルコフ連鎖 time-homogeneous Markov chain という。以下では齊時マルコフ連鎖について議論する。このとき、状態が1期のうちに j から k へ移る確率

$$p_{jk} = P(X_{n+1}=k | X_n=j)$$

を、推移確率 transition probability という。また、 p_{jk} を (j, k) 成分を持つ行列 P を推移確率行列 transition probability matrix という。確率の性質から各行の成分は

$$\sum_{k=1}^N p_{jk} = 1$$

をみたす(この性質は線形代数学では一般に確率行列 stochastic matrix といわれる)。

ここで2つほどマルコフ連鎖の例を挙げよう。最初の例はすでに10.1節で説明したランダム・ウォークと破産問題である。2番目は生物の集団遺伝学への応用例である。

例10.2 <ランダム・ウォークと破産問題> ランダム・ウォークでは、状態は $0, \pm 1, \pm 2, \dots$ であり、推移確率を

$$p_{jk} = P(S_{n+1}=k | S_n=j)$$

とおけば

$$\begin{aligned} p_{jk} &= p && (k=j+1 のとき) \\ &= q && (k=j-1 のとき) \\ &= 0 && (\text{その他}) \end{aligned}$$

である。一方、破産問題では A が n 時点で保持するチップの枚数を X_n とすれば、状態は $0, 1, \dots, N$ で、推移確率は

$$p_{jk} = p \quad (1 \leq j \leq N-1, k=j+1 のとき)$$

$$\begin{aligned} &= q && (1 \leq j \leq N-1, k=j-1 のとき) \\ &= 1 && (j=k=0 \text{ あるいは } j=k=N のとき) \\ &= 0 && (\text{その他}) \end{aligned}$$

となる。ここでもし相手が破産した場合、各々が確率 p あるいは q で相手に一枚チップを施したとする。このとき推移確率は

$$p_{00}=q, \quad p_{01}=p, \quad \text{および} \quad p_{N,N-1}=q, \quad p_{NN}=p$$

となる。これを反射壁 reflecting barrier をもつランダム・ウォークという。

例10.3 <集団遺伝学> 集団の個体数は、各世代において N に固定しておく。遺伝因子としては A と a があり、各個体は2つの遺伝因子を持っている。いま雌雄同株として、遺伝子はいったんプールされたのち、図10.6のようにランダムに復元抽出により交配し、 N 個の次世代の個体が作られる。いま第 n 代において、 a の総数を X_n とすれば、 A の総数 $2N - X_n$ である。このとき状態は $0, 1, 2, \dots, 2N$ からなり、推移確率は

$$p_{jk} = {}_{2N}C_k \left(\frac{j}{2N}\right)^k \left(1 - \frac{j}{2N}\right)^{2N-k}$$

である。これは二項分布と同じ形である。

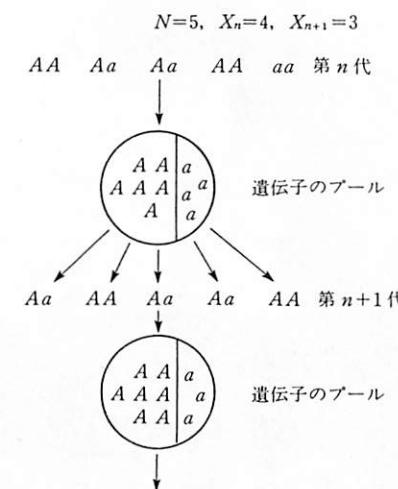


図10.6 生物の集団遺伝の例
マルコフ連鎖が大きな貢献を果たした分野として生物学の集団遺伝がある。雌雄同株の個体群は、対立する遺伝因子 A, a をいったんプールしたのち、任意交配により子孫を残していく。

10.3.2 行列による表現

マルコフ連鎖 $\{X_n\}$ で、2期先以降の状態にたいする推移確率、推移確率行列を考えよう。いま

$$p_{jk}^{(2)} = P(X_{n+2}=k \mid X_n=j)$$

とおく。 X_{n+1} における状態はなんでもよいので

$$\begin{aligned} p_{jk}^{(2)} &= \sum_{l=1}^N P(X_{n+2}=k, X_{n+1}=l \mid X_n=j) \\ &= \sum_{l=1}^N P(X_{n+2}=k \mid X_{n+1}=l, X_n=j) P(X_{n+1}=l \mid X_n=j) \end{aligned}$$

となる。さらに齊時マルコフ連鎖であることに注意すれば

$$\begin{aligned} &= \sum_{l=1}^N P(X_{n+2}=k \mid X_{n+1}=l) P(X_{n+1}=l \mid X_n=j) \\ &= \sum_{l=1}^N p_{lk} p_{lj} \end{aligned}$$

となる。したがって、2期先の推移確率行列は行列の積で $P \cdot P = P^2$ 、推移確率はその各成分である。同様にして n 期先の状態に対する推移確率行列は P^n 、推移確率 $p_{jk}^{(n)}$ はその各成分である。任意の自然数 m, n に対して

$$P^{m+n} = P^m P^n$$

が成立するので

$$p_{jk}^{(m+n)} = \sum_{l=1}^N p_{jl}^{(m)} p_{lk}^{(n)}$$

である。これをチャップマン・コルモゴロフの方程式 Chapman-Kolmogorov equation という。

次に時点 n で、状態 j である確率を求めよう。

$$p_n(j) = P(X_n=j) \quad j=1, 2, \dots, N$$

とおき、ベクトル \mathbf{p}_n を

$$\mathbf{p}_n = (p_n(1), p_n(2), \dots, p_n(N))$$

と定義する。これを状態確率分布という。とくに時点 0 に対しては、 \mathbf{p}_0 を初期分布という。いま

$$\begin{aligned} p_n(j) &= \sum_{l=1}^N P(X_{n-1}=l) P(X_n=j \mid X_{n-1}=l) \\ &= \sum_{l=1}^N p_{n-1}(l) p_{lj} \end{aligned}$$

に注意すれば、ベクトル表現になおして

$$(10.8) \quad \mathbf{p}_n = \mathbf{p}_{n-1} P$$

が成立する。この式を順次くり返せば

$$\mathbf{p}_n = \mathbf{p}_0 P^n$$

を得る。したがって任意の時点における状態確率分布は、初期分布と推移確率行列さえわかれば、決定できる。

さて、時点 n の推移とともに、 \mathbf{p}_n は極限的にどのように変化していくであろうか。もし初期条件に無関係に $\pi = \lim_{n \rightarrow \infty} \mathbf{p}_n$ が存在し、 $\pi = (\pi_1, \pi_2, \dots, \pi_N)$ が確率分布になっているとき、これを定常分布 stationary distribution、あるいは(定常的になればこれ以上は変化しないことから)不变分布 invariant distribution という。一般に、定常分布が存在するか否かの証明は難しい場合がある。しかし存在を仮定すれば、(10.8)の両辺において $n \rightarrow \infty$ として

$$\pi = \pi P$$

を満足する π を求めればよい。

例 10.4 〈反射壁をもつランダム・ウォーク〉 この場合、 $\pi = (\pi_0, \pi_1, \pi_2, \dots, \pi_N)$ とおくと

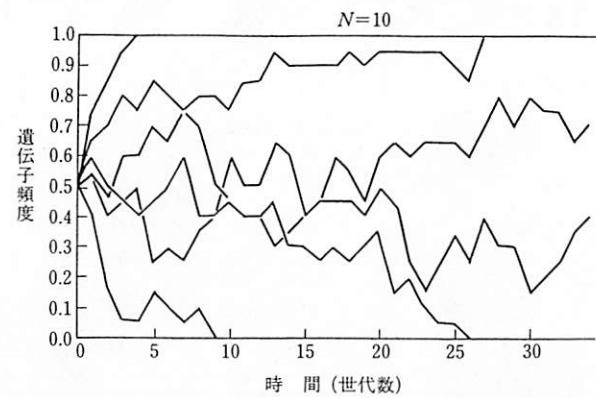
※シミュレーションによる集団遺伝学※

確率分布、確率過程などは複雑になればなるほど、理論的にその性質を解明するのは大変になる。そのとき大きな威力を発揮するのがコンピュータを用いたシミュレーションである。公平なさいころか否かは、何十回、何百回となく投げていれば見当がつくように、擬似乱数を発生させて、標本関数の挙動を探るわけである。(乱数については第11章参照のこと)

図 10.7(a)は集団遺伝学の例である。ここでは遺伝因子 a の第 0 世代における割合(初期頻度) $X_0/2N$ を 0.5 に固定する。マルコフ連鎖としては、初期分布は確率 1 で状態 N となっている。このとき、個体数 N を 10 としたときのさまざまな標本関数 $\{X_n ; n=0, 1, 2, \dots\}$ の軌跡が描かれる。第 4 世代ほどですべて a となることもあれば、第 9 世代ほどで絶滅する可能性もある。われわれが今日、生きているということは非常に幸運なことかも知れない。

なお図 10.7(b)は遺伝因子 a の初期頻度を 0.5、個体数 N を 5 としたときの各世代(1, 5, 10, 15 代)における頻度の確率分布を表している。左側の黒い柱は a が絶滅する確率を、右側の黒い柱はすべて a に固定する確率を示している。世代を経るごとにこの確率は大きくなり、第 15 世代では、各々 30% を超える。

$$\begin{aligned}\pi_0 &= q\pi_0 + q\pi_1 \\ \pi_j &= p\pi_{j-1} + q\pi_{j+1} \quad (j=1, 2, \dots, N-1) \\ \pi_N &= p\pi_{N-1} + p\pi_N\end{aligned}$$



(a) 遺伝子頻度の変化のシミュレーション

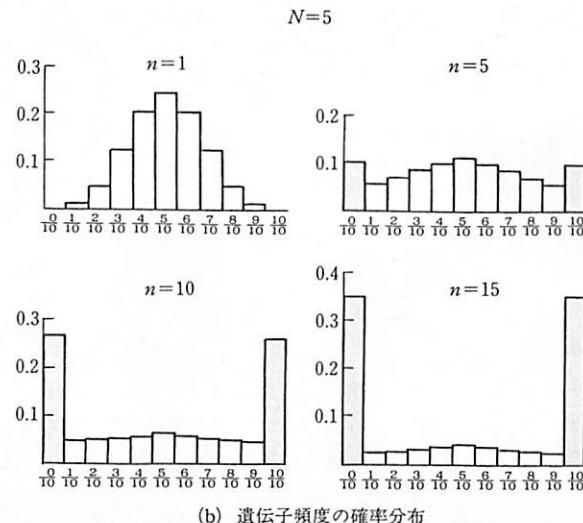


図 10.7 集団遺伝学のシミュレーション例(前頁参照)

理論的に性質を導出するのが困難な確率現象の挙動を解明するのに、大きな威力を発揮するのがコンピュータを用いたシミュレーションである。遺伝子頻度の変化も一目瞭然、把握可能である。

(出典：木村資生『分子進化の中立説』紀伊国屋書店)

となる。したがって

$$p\pi_0 = q\pi_1, \quad p\pi_1 = q\pi_2, \quad \dots, \quad p\pi_{N-1} = q\pi_N$$

であるから

$$\pi_j = (p/q)^j \pi_0 \quad (j=0, 1, \dots, N)$$

を得る。ここで確率であるという条件

$$\sum_{j=0}^N \pi_j = 1$$

より、極限的に近づいていく確率分布は

$$\begin{aligned}\pi_j &= \frac{1 - (p/q)}{1 - (p/q)^{N+1}} (p/q)^j \quad (p \neq q \text{ のとき}) \\ &= 1/(N+1) \quad (p = q \text{ のとき})\end{aligned}$$

となる。したがって、 $p=q=1/2$ であればどの状態にあるのも等確率である。一方、 $p>q$ であれば反射壁 N に近づくほど、 $p<q$ のときには反射壁 0 に近づくほど、確率は大きくなる。いずれも当然の結果であろう。

10.4 ポアソン過程と出生死滅過程

10.4.1 ポアソン過程

連続的に観測する場合のマルコフ性も同様に、 $s < t$ に対して

$$P(X_t = k | X_r = j, r < s, X_s = j) \equiv P(X_t = k | X_s = j)$$

で定義される。この場合のマルコフ連鎖として、まずポアソン過程をとりあげよう。

第1章のポアソン分布は、単位時間内にある事象が起きる回数についての分布である。ここで単位時間に限定せず、任意の時間内 $[0, t]$ にある事象が起きた回数を N_t と表し、確率過程とみなしたのがポアソン過程 Poisson process である。そこでは、二項分布から出発して、事象が生起する区間の幅を細かくしていった極限として、ポアソン分布を導いた。ここでは、より直接的に N_t 自体にいくつかの仮定をおいて、 N_t の状態の確率分布を導こう。

まず N_t は、状態 $0, 1, 2, \dots$ をとる確率過程で、時点 0 では、確率 1 で $N_0 = 0$ とする。また区間が長いほど起きる回数は多くなるから $N_s \leq N_t (s \leq t)$ である。つぎに $0 < s \leq t$ として区間 $[0, s]$ と区間 $[s, t]$ とで起きる回数、つまり

$N_s - N_0$ と $N_t - N_s$ とは独立とする。最後に推移確率に対しては

$$(10.9) \quad P(N_{t+h}=n+1 | N_t=n) = \lambda h + o(h)$$

$$(10.10) \quad P(N_{t+h}=n | N_t=n) = 1 - \lambda h + o(h)$$

を仮定する。ここで一般に $o(h)$ は^{*}、 $h \rightarrow 0$ のとき、 $o(h)/h$ が 0 へ収束する（つまり、 h より‘速く’0 へ近づく）関数を表す数学記号である。ただし(10.9) と(10.10) の $o(h)$ は、同じ関数とは限らない。

独立性より(10.9) と(10.10) は

$$\begin{aligned} P(N_{t+h}=n+1 | N_t=n) &= P(N_{t+h}-N_t=1 | N_t=n) \\ &= P(N_{t+h}-N_t=1) \\ &= \lambda h + o(h) \end{aligned}$$

および

$$\begin{aligned} P(N_{t+h}=n | N_t=n) &= P(N_{t+h}-N_t=0 | N_t=n) \\ &= P(N_{t+h}-N_t=0) \\ &= 1 - \lambda h + o(h) \end{aligned}$$

を意味する。つまり区間 $(t, t+h]$ で起きる回数の確率は、 t に依存せず区間の幅 h のみによって決まる。したがって、この仮定はベルヌーイ試行の定常性に相当する。また(10.9) と(10.10) から、推移確率に対して

$$\begin{aligned} P(N_{t+h} \geq n+2 | N_t=n) &= 1 - P(N_{t+h}=n \text{ あるいは } n+1 | N_t=n) \\ &= 1 - \{\lambda h + o(h)\} - \{1 - \lambda h + o(h)\} \\ &= o(h) \end{aligned}$$

が成立する。つまり区間幅 h が 0 に近づくとき、2 回以上生起する確率は無視でき、微小な区間では状態の変化が緩やかであるとする。この条件はベルヌーイ試行の二値性に対応して、希少性という。以上の条件のもとで

$$(10.11) \quad P(N_t=k) = (\lambda t)^k e^{-\lambda t} / k! \quad (k=0, 1, 2, \dots)$$

が成立する。

証明の概略を示そう。条件付き確率の性質から、 $k \geq 1$ に対して

$$\begin{aligned} (10.12) \quad P(N_{t+h}=k) &= P(N_{t+h}=k | N_t=k-1)P(N_t=k-1) \\ &= \{\lambda h + o(h)\}P(N_t=k-1) + \{1 - \lambda h + o(h)\}P(N_t=k) + o(h) \end{aligned}$$

*) “スモール・オー”と読む。たとえば $h^{3/2}, h^2$ などはこれにあたる。

$$\begin{aligned} &+ P(N_{t+h}=k | N_t=k)P(N_t=k) + o(h) \\ &= \{\lambda h + o(h)\}P(N_t=k-1) + \{1 - \lambda h + o(h)\}P(N_t=k) + o(h) \\ &= \lambda h P(N_t=k-1) + (1 - \lambda h)P(N_t=k) + o(h) \end{aligned}$$

が成立する。ここで

$$p_h(t) = P(N_t=k)$$

とおけば、(10.12) より

$$(10.13) \quad p_h(t+h) - p_h(t) = \lambda h \{p_{h-1}(t) - p_h(t)\} + o(h)$$

となる。そこで(10.13) の両辺を h で割って、 $h \rightarrow 0$ とすると、 $o(h)/h \rightarrow 0$ から

$$(10.14) \quad p'_h(t) = \lambda \{p_{h-1}(t) - p_h(t)\} \quad (k=1, 2, \dots)$$

を得る。 $k=0$ のときも同様に

$$(10.15) \quad p'_0(t) = -\lambda p_0(t)$$

が導ける。(10.14) と(10.15) は、 $p_k(t)$, ($k=0, 1, 2, \dots$) に対する微分差分方程式の体系である。 $N_0=0$ であるから、境界条件

$$\begin{aligned} p_k(0) &= 1 & (k=0), \\ &= 0 & (k \neq 0) \end{aligned}$$

のもとで、帰納法により(10.11) が導ける（詳細は参考文献を参照のこと）。

例 10.5 〈 α 粒子のカウント数〉 放射性物質から飛び出る α 粒子がスクリーンにあたる個数が、ポアソン分布で近似できることは経験的に知られているが、ベイトマン(Bateman, H.)はこれが理論的にも正当化できることを上の方法で示した。実際のポアソン過程の標本関数は、図 10.8 のように、ジャンプする階段関数になることが知られている。ジャンプする時点(図 10.8 では T_1, T_2, \dots)において事象(放出)が起きている。

10.4.2 出生死滅過程

ポアソン過程は事象が起きる度に、それを数え上げていく過程であるから、図 10.8 のように時間と共に増加していく。一方でバクテリア、アメーバなど

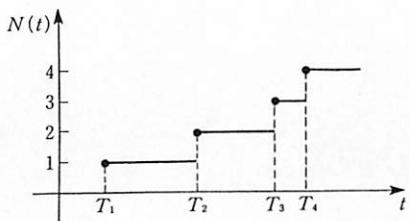


図 10.8 ポアソン過程の挙動

ポアソン過程の標本関数は、事象が起きる時点で 1 段上がる階段関数である。タクシーや料金に似ている。ただし、タクシー料金は段の幅が一定であるが、ポアソン過程では事象が起きる間隔がランダムである。

の微生物、細菌の総数は、分裂したり、子を生むことにより増えることもあるが、死亡により減ることもある。このような増加と減少を共に含む現象を表現するマルコフ連鎖が一般に出生死滅過程である。

いま時点 t におけるある細菌の総数を L_t としよう。総数の増減について推移確率が、仮定

$$P(L_{t+h}=k+1|L_t=k)=\lambda_k h + o(h)$$

$$P(L_{t+h}=k|L_t=k)=1-\lambda_k h - \mu_k h + o(h)$$

$$P(L_{t+h}=k-1|L_t=k)=\mu_k h + o(h)$$

$$P(L_{t+h}\leq k-2 \text{ あるいは } \geq k+2|L_t=k)=o(h)$$

をみたすとき、 L_t を出生死滅過程 birth and death process という。総個体数が k のとき、区間 $(t, t+h]$ において 1 つの個体が生まれる確率が

$$\lambda_k h + o(h)$$

であり、死滅する確率が

$$\mu_k h + o(h)$$

である。ポアソン過程との違いは、総数が減る確率が 0 とは限らないこと、また増減の確率が現在の総数 k に依存することである。したがって、独立増分過程ではない。ここで

$$p_k(t)=P(L_t=k)$$

とおこう。ポアソン過程の(10.14)を導いたのと同様の方法で、 $k \geq 1$ に対して

$$(10.16) \quad p'_k(t)=-(\lambda_k + \mu_k)p_k(t) + \lambda_{k-1}p_{k-1}(t) + \mu_{k+1}p_{k+1}(t),$$

$k=0$ に対しては、0 から減ることはないので

$$(10.17) \quad p'_0(t)=-\lambda_0 p_0(t) + \mu_1 p_1(t)$$

が導ける。

λ_k および μ_k の値を変えることによりさまざまな確率過程を作り出すことができる。たとえば

$$\lambda_k \equiv \lambda, \quad \mu_k \equiv 0$$

とおけば出生死滅過程はポアソン過程になる。たしかに(10.16)、(10.17)はそれぞれ(10.14)、(10.15)に帰着する。また λ_k, μ_k を個体数に比例させ

$$(10.18) \quad \lambda_k = k\lambda, \quad \mu_k = k\mu$$

とおいたものを、フェラー・アレイ過程 Feller-Arley process という。フェ

ラー・アレイ過程で、死滅を考えず

$$\mu_k = 0$$

としたとき、純出生過程 pure birth process あるいはユール過程 Yule process という。さらにはフェラー・アレイ過程で外部からの移民 immigration を考え、区間 $(t, t+h]$ の間に新たな個体が移住してくる確率を

$$\nu h + o(h)$$

とおいて

$$\lambda_k = k\lambda + \nu$$

としたとき、ケンドール過程 Kendall process という。

ポアソン過程に対する(10.15)と違って、(10.17)の右辺には $p_1(t)$ が含まれているので、出生死滅過程の $p_k(t)$ を帰納的に導くことはできない。個別的に解を求めるしかないが、とくにフェラー・アレイ過程については、

$$(10.19) \quad p_0(t) = \frac{\mu \{1 - \exp((\lambda - \mu)t)\}}{\mu - \lambda \exp((\lambda - \mu)t)} \quad (\mu \neq \lambda \text{ のとき})$$

$$(10.20) \quad = \lambda t / (1 + \lambda t) \quad (\mu = \lambda \text{ のとき})$$

$k \geq 1$ に対しては

$$p_k(t) = \frac{\lambda^{k-1}(\lambda - \mu)^2 e^{(\lambda - \mu)t} (1 - e^{(\lambda - \mu)t})^{k-1}}{(\mu - \lambda e^{(\lambda - \mu)t})^{k+1}} \quad (\mu \neq \lambda \text{ のとき})$$

$$= (\lambda t)^{k-1} / (1 + \lambda t)^{k+1} \quad (\mu = \lambda \text{ のとき})$$

が示されている。ただし、境界条件を

$$p_1(0) = 1, \quad p_k(0) = 0 \quad (k \neq 1)$$

とする。

例 10.6 〈出生率と死亡率〉 フェラー・アレイ過程について、定常分布 $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ を求めよう。(10.16), (10.17) に(10.18)を代入して $t \rightarrow \infty$ とする。このとき $p'_k(t)$ の極限値は 0 とし、 $\lambda_0 = 0$ に注意すれば

$$\mu \pi_1 = 0$$

$$(\lambda + \mu)k\pi_k = \lambda(k-1)\pi_{k-1} + \mu(k+1)\pi_{k+1} \quad (k \geq 1 \text{ のとき})$$

が成立する。したがって

$$\pi_1 = \pi_2 = \pi_3 = \dots = 0$$

である。もし $\pi_0 = 1$ であれば極限ではすべて死滅することになる。一方、 $\pi_0 <$

1であれば $1 - \pi_0$ の確率で個体数は「爆発」することになる。どちらになるかはひとえに $p_0(t)$ の極限の値に依存する。(10.19), (10.20) より

$$\begin{aligned}\pi_0 &= 1 & (\lambda \leq \mu) \\ &= \mu/\lambda & (\lambda > \mu)\end{aligned}$$

である。確かに $\lambda < \mu$ で出生率が死亡率を下回れば、最終的にはすべて死滅してしまうことは想像できる。しかし、 $\lambda = \mu$ で出生率と死亡率が等しくても同じ結果になることはやや意外である。このことは昨今の人口問題を考える際の参考になるであろう。

10.5 確率過程の応用例

10.5.1 待ち時間分布と待ち行列

待ち時間分布 いまポアソン過程で、時点 0 から出発して、最初に事象が起きるまでの時間間隔を T_1 、事象が $(i-1)$ 回起きてから i 回起きるまでの時間間隔を T_i ($i=2, 3, \dots$) とおこう。このとき事象 $\{T_i \leq t\}$ は $\{N_i \geq 1\}$ に等しいので、 T_1 の分布関数を $F_1(t)$ とすれば

$$\begin{aligned}F_1(t) &= P\{T_1 \leq t\} = P\{N_1 \geq 1\} \\ &= \sum_{i=1}^{\infty} (\lambda t)^i e^{-\lambda t} / i! \\ &= 1 - e^{-\lambda t}\end{aligned}$$

となり、その密度関数 $f_1(t)$ は

$$f_1(t) = \lambda e^{-\lambda t} \quad (t > 0)$$

である。つまり、 T_1 の分布は指数分布になっている。これは第1章の指数分布の導き方からみても明らかである。つぎに

$$S_n = T_1 + T_2 + \dots + T_n$$

の分布を考えよう。 S_n は第 n 回目の事象が起きるまでに要する時間である。ポアソン過程は、定常性の仮定を満たし、各増分は互いに独立であるから、 T_i ($i=1, \dots, n$) は互いに独立で、指数分布 $Ex(\lambda)$ に従っている。事象 $\{S_n \leq t\}$ は事象 $\{N_t \geq n\}$ に等しいから、 S_n の確率分布を $F_n(t)$ とすれば

$$\begin{aligned}F_n(t) &= P(S_n \leq t) \\ &= P(N_t \geq n)\end{aligned}$$

$$= \sum_{i=n}^{\infty} (\lambda t)^i e^{-\lambda t} / i!$$

である。密度関数 $f_n(t)$ は、 $F_n(t)$ の右辺の各項を微分して

$$\begin{aligned}f_n(t) &= \sum_{i=n}^{\infty} \lambda \{(\lambda t)^{i-1} e^{-\lambda t} / (i-1)! - (\lambda t)^i e^{-\lambda t} / i!\} \\ &= \lambda \sum_{i=n-1}^{\infty} (\lambda t)^i e^{-\lambda t} / i! - \lambda \sum_{i=n}^{\infty} (\lambda t)^i e^{-\lambda t} / i! \\ &= \lambda (\lambda t)^{n-1} e^{-\lambda t} / (n-1)!\end{aligned}$$

となる(もちろん厳密には、関数の無限和と微分の交換の正当性を示さなくてはいけない)。これは相 n のアーラン分布 Erlang distribution とよばれているが、実は第1章で説明したガンマ分布 $Ga(n, \lambda)$ となっている(問題1.9 参照)。

指数分布を導入したときの例で言えば、相 n のアーラン分布は、前売り乗車券を購入するために並んだ n 番目の人人が、発売開始から購入し終わるまでの待ち時間 waiting time の分布である。アーラン分布は、さまざまなサービスの待ち時間を表す分布として広く利用されている。

待ち行列 この待ち時間の分布をさらに詳しく考えてみよう。前売り乗車券の例でいえば、実際には発売開始後も人々はやってきて列に並び、順番待ちのひとの人数は増加する。一方、駅員はコンピュータに入々の希望を入力して乗車券を発行していく。乗車券を受け取ったひとは立ち去って行くので、このときは列に残っているひとの人数は減る。したがって時間の経過とともに、行列内にいるひとの人数は増えたり減ったりするから、出生死滅過程と同じ状況になっている。人の到着が出生にあたり、サービスの終了が死滅にあたる。このような列(仮定に応じて一般化される)を一般に待ち行列 queue^{*)} という。

いま人々の到着がパラメータ λ のポアソン分布に従い、駅員の処理時間がパラメータ μ (前の節の λ に相当することに注意)の指数分布に従うとすれば、列内にいる人の数は、 $\lambda_k = \lambda$, $\mu_k = \mu$ の出生死滅過程になる。前売り開始時点ですでに i 人の人が並んでいたとすれば、初期分布を

$$\begin{aligned}p_k(0) &= 1 & (k=i) \\ &= 0 & (k \neq i)\end{aligned}$$

とおく。もし駅員が s 人いるばあいには

$$(10.21) \quad \mu_k = k\mu \quad (0 \leq k < s)$$

$$(10.22) \quad = s\mu \quad (k \geq s)$$

*) 発音がむずかしいが、「キュー」と読む。

と変えればよい。ただし人々は1列の順番待ちで、待っている人がいるのに手の空いている係員が出るようなことはないと仮定する。

例10.7 〈待ち行列の長さの定常分布〉 $s=1$ に対する定常分布を求めてみよう。10.4節と同様に(10.21), (10.22)を(10.16), (10.17)に代入して $t \rightarrow \infty$ とすれば、

$$\begin{aligned}\mu\pi_1 - \lambda\pi_0 &= 0 \\ (\lambda + \mu)\pi_k &= \lambda\pi_{k-1} + \mu\pi_{k+1} \quad (k \geq 1)\end{aligned}$$

となる。ここで

$$\rho = \lambda/\mu$$

とおけば、数学的帰納法により

$$\pi_k = \rho^k \pi_0$$

である。したがって $\sum_{k=0}^{\infty} \pi_k = 1$ が成立するためには、

$$\rho < 1, \pi_0 = 1 - \rho$$

でなくてはならない。このとき

$$\pi_k = (1 - \rho) \rho^k$$

である。一方、 $\rho \geq 1$ のときは

$$\sum_{k=0}^{\infty} \rho^k = \infty$$

となるので、 $\pi_0 = 0$ を満足しなくてはならない。このときすべての有限な k に対して $\pi_k = 0$ となるので、時間の推移とともに列に並んでいる人の数(長さ)は発散していくことになる。

ここで10.4節で求めたフェラー・アレイ過程の定常分布との違いに注意しよう。フェラー・アレイ過程では、いったん個体数が0となりすべてが死滅するともはや増えることがない。したがって、 $\lambda = \mu$ のときでも $\pi_0 = 1$ であった。一方、ここでの例では列に並んでる人がある時点でいなくなってしまっても、また新たにほかの人がやってくる可能性がある。そこで $\lambda = \mu$ ($\rho = 1$) のとき、定常分布は発散してしまう。

10.5.2 株価

株価など金融時系列の変動を表現する確率過程としてどのようなものがよいだろうか。今日においても「株価(通常は対数をとる場合が多い)がブラウン運動に従う」という仮説(ブラウン運動仮説)が、実際のデータの解析から支持されるか否かは、重要問題である。

始まりは古く、フランス人バシェリエー(Bachelier, L.)と言われている。彼はアイシュタインに先立つこと5年、1900年にブラウン運動に似たモデルを株価データに応用した論文を発表している。

いま $\{Z_t\}$ をある特定の銘柄の株価あるいは株価指数を表現する確率過程とする。このとき株価収益率を

$$X_t = \log Z_t - \log Z_{t-1}$$

で定義する。ただし、この定義は連続複利とよばれる収益率である。通常収益率というと、

$$X'_t = (Z_t - Z_{t-1})/Z_{t-1}$$

を思い浮かべるが、これは単利の収益率という。現実のデータでは X'_t が小さく、テーラー展開により近似的に

$$X_t \equiv \log(1 + X'_t) \approx X'_t$$

が成立する。したがって、どちらの定義でも本質的な違いはないので、ここでは X_t を採用する。

さて、もし時点 t を連続的に変化させたとき、 $\{\log Z_t\}$ がブラウン運動であるとすれば、離散時点 $t = 0, 1, 2, \dots, T$ で観測した、 $\{X_t\}$ の確率分布は互いに独立で期待値、分散とも一定の正規分布になる。したがって、実際データで検証すべきことは、正規分布に従うか否かと互いに独立で同一の分布に従うか否かの2つの問題に大別される。

ここでは独立性の検定について取り上げよう。この検定に用いられる統計量としては、標本自己相関係数がある。ラグ(遅れ) h の標本自己相関係数 r_h sample autocorrelation coefficient r_h は、

$$r_h = \frac{\sum_{t=1}^{T-h} (X_t - \bar{X})(X_{t+h} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2}$$

で定義される。ここで $\bar{X} = \sum_{t=1}^T X_t / T$ とする。もし $\{X_t\}$ が互いに独立で同一分布に従うと仮定すれば、 $T \rightarrow \infty$ のとき、 r_h ($h = 1, 2, \dots$) は $\rho_h \equiv \text{Cov}(X_t, X_{t+h}) / V(X_t) = 0$ に収束し、 h が異なると互いに独立で正規分布 $N(0, 1/T)$ に近づくことが知られている。したがって現実のデータから計算された r_h が、この事実に反するような動きを示せば、「独立性」の仮説は棄却できる。

例10.8 〈株価データのブラウン運動仮説の検証〉 表10.2は日本の個別銘柄204種の、1983年1月から1987年12月までの5年間の日次データ(約1400個)の、標本自己相関係数を6つのクラス

- (1) $r_h < -0.1$ (2) $-0.1 \leq r_h < -0.05$ (3) $-0.05 \leq r_h < 0.0$
 (4) $0.0 \leq r_h < 0.05$ (5) $0.05 \leq r_h < 0.1$ (6) $0.1 < r_h$

に分類したものである。 h は1から30まで変化させたので、総数は $204 \times 30 = 6120$ になる。正規分布に従うと見なせば標準偏差の2倍は約 $2/\sqrt{1400} = 0.052$ となり、およそ90%については、仮説 $\rho_h = 0$ を棄却できない。だから直ちに独立と判断することはできない。なぜならもし $\{X_t\}$ が互いに独立で同一分布に従う標本列であるならば、絶対値の系列、二乗の系列、 $\{|X_t|\}$, $\{X_t^2\}$ も同じ性質を満たさなければならない。この2つの系列の標本自己相関係数を示したのが、表10.3である。標準偏差の2倍を超え、しかも正の値をとる割合が、表10.2より増えることがわかる。したがって、独立性の成立が疑わしい銘柄数が多くなる。図10.9はある特定の銘柄についてのものだが、同じ特徴が現れている。

$\{|X_t|\}, \{X_t^2\}$ の自己相関係数に正の値が多いのは、株価の大きな下落のあとに反騰、あるいはその逆が生じることによる。原系列では両者が相殺され、自己相関係数は0に近い値となる。また、以上の議論では分散一定を暗黙に仮定しているが、独立な系列であっても分散が一定でなければ、標本自己相関係数の挙動も変わってくる。

表10.2 1~30次のラグの自己相関係数の度数分布

クラス	< -0.1	$-0.1 \sim -0.05$	$-0.05 \sim 0.0$	$0.0 \sim 0.05$	$0.05 \sim 1.0$	> 1.0	計
度数	53	435	2846	2537	245	4	6120

(出典：刈屋・佃・丸編著『日本の株価変動』より)

0の周辺が多いが、それでも独立と判断することはできない。

表10.3 1~30次のラグの自己相関係数の度数分布

(絶対収益率・2乗収益率)

クラス	< -0.1	$-0.1 \sim -0.05$	$-0.05 \sim 0.0$	$0.0 \sim 0.05$	$0.05 \sim 1.0$	> 1.0	計
度数： $ X_t $	0	2	236	1539	2138	2205	6120
度数： X_t^2	0	0	824	3258	1378	660	6120

(出典：同上)

これによっても独立性は確認されず、むしろ疑わしくなる。

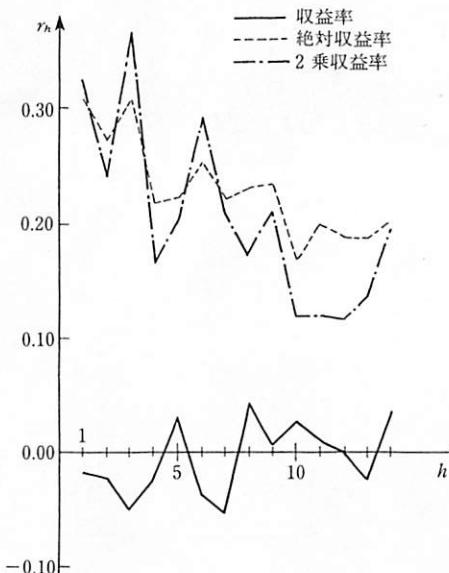


図10.9 ある個別銘柄の自己相関係数
確率変数列の独立性を検定するためには、標本自己相関係数がある。もし独立であれば、各ラグの標本自己相関係数の分布は、標本数が増えると、期待値0分散が標本数の逆数に等しい互いに独立な正規分布に近づく。

また、独立ならば確率変数列は、その絶対値、二乗も互いに独立となるべきであるから、これらの標本自己相関係数の挙動も調べる必要がある。

(出典：前掲)

以上のことから必ずしもすべての銘柄の株価がブラウン運動に従うと断定することはできない。現在までのところ株価に対する決定的なモデルはなく、さまざまな角度から解析が行われている。

付節 差分方程式の解法

数列 $\{a_n\}$ に対する項間の漸化式

$$(1) \quad a_{n+2} + pa_{n+1} + qa_n = 0, \quad n=0, 1, \dots$$

の一般項 a_n を求めることを方程式を解くことと考えた場合、これを(2次または2階の)差分方程式 difference equation という。この解法について説明する。(1)の解が

$$(2) \quad a_n = t^n \quad (t \text{ は定数})$$

の形と仮定して、これを(1)に代入すれば、 t は2次方程式

$$(3) \quad t^2 + pt + q = 0$$

の解でなければならない。これを、差分方程式(1)の特性方程式 characteristic equation という。この2解を α, β とすれば、 $\alpha \neq \beta$ なら

$$(4) \quad a_n = \alpha^n, \quad a_n = \beta^n$$

は(1)の解となる。これを1次独立な解という。さらに一般的なものは、 A, B を任意の定数として

$$(5) \quad a_n = A\alpha^n + B\beta^n$$

となるが、これを(1)の一般解 general solution という。 A, B は、たとえば a_0, a_1 (数列の初項など) が与えられていれば、決定される。

特性方程式の解が重解 α のときは、特別の扱いが必要となる。すなわちこの場合は、容易に確かめられるように $a_n = \alpha^n$ のほかに $a_n = n\alpha^n$ も(1)の解となるので

$$(6) \quad a_n = (A + Bn)\alpha^n, \quad n=0, 1, \dots$$

が一般解となる。

なお、2解(α, β)が複素数であるときは、(5)のままでは扱いにくいので、その絶対値を ρ 、偏角を θ として、指数表示して

$$(7) \quad \alpha = \rho e^{i\theta}, \quad \beta = \rho e^{-i\theta}$$

とすれば、解として(5)はより具体的に

$$(8) \quad a_n = \rho^n (A e^{in\theta} + B e^{-in\theta})$$

$$(8') \quad = \rho^n (A' \sin n\theta + B' \cos n\theta)$$

と表現される。(8')ではド・モアブルの公式を用いた。

〈練習〉 次の差分方程式を解け。

i) $a_{n+2} = a_{n+1} + 2a_n, \quad a_0 = 0, a_1 = 1$

ii) $a_{n+2} = 4a_{n+1} - 4a_n, \quad a_0 = 2, a_1 = 1$

iii) $a_{n+2} = a_{n+1} - a_n, \quad a_0 = 0, a_1 = 1$

iv) $a_{n+2} = a_{n+1} + a_n, \quad a_0 = 1, a_1 = 1$

(フィボナッチ数列)

~~~~~ 第10章 練習問題 ~~~~

10.1 <2次元のランダム・ウォーク> ある粒子は2次元のランダム・ウォークに従い、整数の座標上

$$\{(i, j) : i, j = \dots, -1, 0, 1, \dots\}$$

を動いているとする。 p, q, r, s は $0 < p, q, r, s < 1, p + q + r + s = 1$ をみたす実数とする。もし時点 n において粒子が (i, j) にいたとき、時点 $n+1$ には

$$(i+1, j), \quad (i, j+1), \quad (i-1, j), \quad (i, j-1)$$

に各々確率 p, q, r, s で移動する。また時点毎の移動は互いに独立とする。時点 0 において原点にいたとして、時点 n において原点に再び戻る確率を求めよ。

10.2 <静止を含むランダム・ウォーク> 1次元のランダム・ウォークを考える。粒子は直線上を左に確率 α 、右に確率 β で動き、確率 γ で元の位置に留まっているとする。ここで $0 < \alpha, \beta, \gamma < 1, \alpha + \beta + \gamma = 1$ とする。また原点 0 と N を吸収壁とする。点 a ($0 \leq a \leq N$) から出発したとして 0 に吸収される確率は、

$$p = 1 - q = \beta / (\beta + \alpha)$$

とおけば、(10.2), (10.3) で与えられることを示せ。

10.3 <行動科学的モデル> 4個の箱がある。ひとつの箱には食物が入っている。残りの3つの箱は触ると、電気ショックを感じる。いま3匹のねずみが食物にありつくため箱を開けようとしている。

‘愚かな’ねずみ：毎回でたらめに各箱を開けようとする。

‘普通の’ねずみ：直前に試みた箱は除いて、残り3つの箱をでたらめに開けようとする。

‘利口な’ねずみ：過去に試みた箱はすべて除いて、残りの箱をでたらめに開けようとする。

i) 4個の箱に、1, 2, 3, 4と番号を付ける。番号1の箱に食物が入っている。普通のねずみが n 回目に開ける箱を X_n とする。 X_n の推移確率を求めよ。ただし、食物の入ってる箱がわかった後は終生忘れないとする。

ii) 各ねずみが食物にありつくまでの試行回数の期待値を求めよ。

10.4 <平均個体数> フェラー・アレイ過程での時点 t における平均個体数を

$$m(t) = \sum_{k=1}^{\infty} kp_k(t)$$

とおく。このとき

$$m'(t) = (\lambda - \mu)m(t)$$

を示せ。また初期条件

$$p_i(0)=1, \quad p_k(0)=0 \quad (k \neq i)$$

のもとで、 $m(t)$ を求めよ。

10.5 <ユール過程の解> いま、 $p_1(0)=1, p_k(0)=0 \quad (k \neq 1)$ とし、さらに
 $p_k(t)=e^{-\lambda t}(1-e^{-\lambda t})^{k-1} \quad (t > 0, k \geq 1)$

とする。この解はユール過程に対する微分方程式を満足することを確かめよ。ただし、ユール過程においては $\lambda_0=\mu_0=0$ となるので、 $k \geq 1$ に対する微分方程式を考えればよい。

第11章 亂数の性質

乱数とは、「ランダム・ナンバー」、すなわちランダムに並んだ数という意味である。ランダムであるためにはどのような条件を満たせばよいだろうか。また、シミュレーションに用いるには、どのようにして発生させればよいであろうか。統計学に関係の深いこの乱数について、基礎的なことがらをまとめておこう。

11.1 亂数の性質

われわれがよく耳にする「乱数」ということばは、次のような場面に結びついている。

- a) プロ野球でバッテリーが交換するサイン
- b) テレビ・ニュースで報道された逮捕されたスパイ容疑者が持っていた乱数表
- c) 統計学の講義でふれられた「シミュレーション」や「モンテ・カルロ法」で用いられる乱数

この乱数 random number や乱数表 table of random numbers はどのようにして作られるのであろうか。どのような性質をもっている(あるいは、もってべき)ものであるのか、それを解説しよう。

まず、場面 a), b) で使われている乱数表がもっているべき性質は同一であり、

表11.1 亂数表

82 69 41 01 98	53 38 38 77 96	38 21 08 78 41	21 91 44 58 34	29 73 80 76 80
17 66 04 63 41	77 51 83 33 14	04 23 86 16 23	44 37 81 32 71	14 62 21 91 11
58 26 41 01 59	68 98 40 57 93	41 58 15 53 52	48 67 96 77 09	40 04 65 63 09
07 16 73 31 65	61 64 17 83 92	67 70 62 34 65	61 85 15 24 36	19 72 16 57 20
13 43 40 20 44	75 93 89 23 44	59 95 05 42 31	89 35 88 85 65	23 85 04 45 44
26 86 01 11 93	19 96 29 40 36	03 99 67 87 54	25 16 38 69 73	05 31 83 78 55
38 75 35 82 11	00 81 89 17 75	55 50 22 45 74	66 78 10 03 70	95 31 91 23 99
62 86 84 47 47	44 88 10 83 73	68 40 94 81 56	91 80 40 87 71	79 78 05 23 45
62 88 58 97 83	35 14 27 88 69	38 03 25 20 18	98 84 74 10 38	08 81 57 44 38
56 63 41 73 69	71 11 08 02 22	54 93 82 38 95	39 87 63 52 59	84 32 98 57 87
22 64 95 98 05	66 83 86 98 01	11 47 12 32 05	46 72 06 63 42	12 91 15 16 02
92 77 38 93 35	66 98 43 50 87	12 93 49 62 27	91 05 93 32 41	64 70 29 43 23
72 31 02 74 28	95 57 25 71 05	93 87 29 72 20	44 98 06 84 76	70 63 94 37 87
96 24 11 47 32	79 92 28 60 76	98 90 99 07 13	21 96 72 64 40	04 32 19 67 44
92 76 08 37 03	42 02 88 03 51	61 82 12 67 26	18 14 34 71 53	42 21 15 30 81
62 30 11 43 58	64 54 72 13 14	15 17 41 35 56	52 55 38 84 74	32 38 80 07 15
44 80 77 97 30	33 80 68 83 88	11 60 03 40 54	35 00 27 70 53	61 81 42 23 70
47 07 55 30 25	42 02 47 27 28	29 74 00 44 69	96 60 99 05 62	23 66 61 11 03
64 07 82 05 27	12 84 96 51 74	05 20 84 45 66	11 83 71 31 78	62 39 47 74 65
83 65 17 55 22	11 24 61 41 94	25 24 88 49 42	29 33 60 93 50	39 95 37 96 97
56 78 68 91 56	20 25 96 71 17	23 72 44 40 33	00 60 14 29 15	41 33 48 22 60
94 14 77 00 08	03 31 01 74 18	01 94 77 17 39	68 95 26 16 33	29 91 83 85 63
89 03 76 89 00	66 41 72 40 99	30 79 17 58 52	28 69 56 31 05	10 67 40 89 04
47 50 75 77 58	50 10 81 87 28	78 97 60 70 74	84 38 89 42 22	53 95 41 39 64
13 53 30 19 65	45 70 06 41 99	38 90 71 38 65	16 03 27 39 54	44 48 62 81 42
56 15 07 26 23	03 20 27 68 53	52 23 56 99 08	38 16 66 94 90	93 27 29 85 52
66 57 11 72 47	49 99 75 81 49	11 33 01 53 46	48 84 62 51 07	38 48 37 84 61
18 70 75 69 83	27 42 08 42 32	98 09 18 30 08	50 43 68 29 16	41 52 51 74 18
80 01 74 84 64	85 60 18 90 05	04 89 02 21 99	66 08 34 08 51	76 98 69 45 68
73 09 21 10 26	42 76 96 96 67	38 31 80 14 95	85 24 21 21 98	59 64 81 65 55
51 10 26 95 56	14 57 33 37 48	40 89 46 24 36	96 76 09 00 19	69 54 06 09 53
76 86 54 99 70	94 22 80 66 42	98 99 68 17 57	58 82 15 79 48	03 57 64 62 35
59 58 40 46 54	75 46 74 70 53	27 08 91 73 59	38 40 46 81 13	68 45 90 02 87
76 78 86 82 37	92 71 64 35 88	73 84 41 37 88	64 95 23 72 03	79 91 71 30 04
80 58 54 62 80	94 10 14 54 26	86 37 72 29 78	13 56 65 62 38	56 59 90 27 29
69 30 74 71 17	02 37 55 92 73	33 14 21 87 08	12 77 97 29 42	94 47 82 27 22
08 20 69 34 34	60 92 83 45 49	66 38 31 51 48	57 02 11 40 22	15 25 88 06 57
37 80 59 15 14	30 44 06 91 66	00 77 11 19 38	14 84 97 82 26	45 14 85 99 20
81 45 72 59 90	57 50 22 04 27	53 23 00 49 15	49 27 83 13 33	93 64 66 36 77
23 03 76 70 82	29 35 94 85 13	68 46 89 22 66	24 01 96 27 73	96 00 88 65 16
73 39 18 51 24	23 89 51 91 16	26 52 05 39 87	61 49 26 75 81	35 89 21 99 48
91 28 53 00 70	16 18 39 81 82	09 86 94 36 59	17 15 51 37 23	68 19 64 93 74
46 82 06 04 38	20 67 31 59 26	39 73 23 24 24	14 06 87 09 13	00 30 38 38 05
81 15 86 25 86	07 58 60 18 93	52 52 04 59 53	61 82 17 08 81	91 90 66 67 39
43 77 34 49 86	98 20 99 18 81	92 46 75 32 82	84 60 96 09 60	57 26 23 36 11
67 26 67 28 42	03 30 79 21 30	73 85 83 99 12	42 12 89 70 86	46 25 58 00 80
38 25 56 68 83	92 02 54 80 38	51 66 56 77 51	75 48 11 37 18	79 81 36 25 93
90 15 30 77 30	47 72 29 66 14	49 96 25 45 96	51 40 71 47 49	63 43 30 38 92
61 53 95 73 24	87 94 87 35 18	59 18 82 99 75	80 55 80 89 73	10 74 47 86 85
85 59 27 83 53	19 80 44 68 77	86 86 73 88 75	98 06 98 65 01	77 78 86 79 60

(出典：森口繁一(編)『新編日科技連数值表』、日科技連出版社、1990年)

11.1 亂数の性質

それは次のようなものであると考えられる。

数字の並び方に簡単な規則性がなく、したがって乱数表の一部分を他人に見られたとしても、見られなかった部分にどのような数字の列がならんでいるか断定される心配がない。この性質は予測不能性 unpredictability とよばれる。

これに対して、c)の例で使われる乱数は必ずしもこのような性質を備えてい必要はなく、むしろ簡単な規則に従って作られたものが使われる。

ここまで、単に「乱数」あるいは「乱数表」という言葉を使ってきたが、これらを2種類——一様乱数(表)とその他の乱数(表)——に分けて考えるのが便利である。一様乱数表といえば、ふつう表11.1のように0から9までの数字が長方形状(行列のかたち)に並んでいるものをさす。表で、適当にスペースが入れてあるのは、見やすくするためである。このようなものを作るのは原理的には簡単であり、次のようにすればよい。まず碁石を10個用意し、これらに0から9までの数字を、1個の碁石に1個ずつ、かつ書かれぬ数字がないよう、もれなく記入した後、不透明な袋の中に入れる(碁石がなければ、形と大きさが同じガラス玉などを使ってもよい)。以上の準備の後、次の操作を繰り返すことによって一様乱数の数列を得る：袋の中の碁石をよくかきませてから1個を取り出し、それに書いてある数字を記入してから碁石を袋にもどす。こうして得られた数字を順番に、左から右へ、上から下へ、並べていけば、一様乱数表が得られる。

なお、以上の操作をもっと簡単にできるようにするために、「乱数サイ」と呼ばれるサイコロも市販されている。これは、正二十面体の各面に0から9までの数字のいずれかひとつが書いてあり、しかもどの数字もちょうど二つの面上に書いてあるものである。

「確率変数」ということば^{*)}を使っていえば、0から9までの整数值を等確率(すなわち各整数值を1/10の確率)でとる確率変数Xの実現値(観測値)の系列を記録したものが一様乱数表である。したがって、各数字の出現度数を乱数表全体にわたって調べてみれば、それらはほぼ一様である(一様分布に近い)と期待される。また、一様乱数表を‘10進2けたの数字の集まり’と見なして、その度数分布を調べてみても、やはり一様分布に近いと期待されることも容易に

*) 本シリーズI『統計学入門』第5章

理解できる。3けた以上の場合についても同様である。

実験が小規模で、使う乱数がそれほど多くない場合には、以上のような操作を自分で行って乱数を作り出してもよいし、あるいは市販されている乱数表を使ってもよいが、大量に乱数を必要とする場合には不便である。大量に乱数を使う実験はコンピュータを使って行うのがふつうなので、乱数もコンピュータの中で作り出せないかと考えるのは自然である。

実際、コンピュータの創案者であるフォン・ノイマン(von Neumann, J., 1903.12-1957.2)自身も、このようなことを考え、ある方法を提案している(ただし、この方

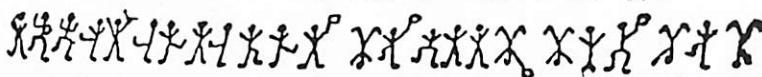
※名探偵による統計学講義※

名探偵シャーロック・ホームズ Sherlock Holmes で知られるコナン・ドイル(C. Doyle, 1859-1930)の一短篇『踊る人間』The Dancing Men は、暗号から題材をとったものとしてよく知られている。

主人公(依頼人)は、ロンドンの東北、美しい田園地帯ノーザン・ノーフォーク Norfolk (州)に代々 5 世紀も続いた莊園主の家系の当主だが、その貞実なアメリカ人の妻が過去の生い立ちのゆえに、暗闇の奥の何者かによって人形文字の暗号で脅迫されるという筋書き。話は、リドリング・ソープ Riddling Thorpe ——日本風には「なぞかけ村」となる——という架空^{*)}の小村で一枚の紙切れが拾われることから始まる。(以下、岩波文庫『シャーロック・ホームズの帰還』より)

「もしこれが全然でたらめなものなら、これを解くことは不可能かもしれません。が、反対に組織的(systematic)なものなら、きっと底をつきとめることができます。しかし、このここにある見本(sample)は、これはあまり短すぎて、どうにもなりません。」名探偵による統計学の講義である。

ホームズは、英語のアルファベットの生起頻度が E, T, A, O, I, N, S, H, R, D, L, … の順だとして、この暗号を面白いように解く。その一例：



ELSIE PREPARE TO MEET THY GOOD
(解読結果)

メッセージは「エルジーよ、汝の神に会う用意をせよ」(覚悟せよ)を意味する。推理小説なので結末を言うのはやめよう。

^{*)} まったくの架空ではなく、実在地名を組み合わせてある。

法は現在では使われていない)。

コンピュータの中での乱数の発生は、あらかじめ用意されているプログラムに従って行われるのがふつうである。このようにして発生される乱数を上で述べた意味での乱数と区別するために算術乱数あるいは擬似乱数とよぶことがある。算術乱数は、乱数が本来もっているべき予測不能性という性質を持たないので、暗号のために使うのは危険であるが、一般的なシミュレーションではこの性質は不要であるので、特にさしつかえなく使えることが多い。

以下で算術乱数の発生方法について述べよう。なお、今後は算術乱数のことを単に乱数とよぶことにする。

11.2 一様乱数の発生法

11.2.1 線形合同法

1948 年頃にレーマー(Lehmer, D. H.)が提案した方法で、広く使われている。あとで述べるような欠点があるので、大量の乱数を使って精度の高いシミュレーション結果を得ようとする実験には向かないが、小規模な実験用には簡単で便利な乱数発生法である。

市販されているパソコンのプログラム言語 Basic や Pascal 等、また大型コンピュータやワークステーション上の UNIX^{*)} システムで標準的に用意されている乱数発生機能は、たいていこの方法に基づいている。

この線形合同法 linear congruential method は漸化式

$$(11.1) \quad X_n = aX_{n-1} + c \pmod{M}, \quad n \geq 1$$

を使って乱数列 X_0, X_1, X_2, \dots を出力する。ここで乗数 a および M は、以下に述べる基準にしたがって選ぶべき正の整数であり、加数 c は 0 または正の整数である。初期値 X_0 を適当に選んで、 $n=1$ として(11.1)を使うと

$$X_1 = aX_0 + c \pmod{M}$$

^{*)} UNIX はベル研究所(AT & T)の登録商標である。

となる。すなわち、 $aX_0 + c$ を M で割った余り(剰余)を X_1 とする。この M を法 modulus という^{*}。 X_2 が必要になったら、こんどは(11.1)式を $n=2$ として使う。以下同様である。すなわち、新しい乱数が必要になったら、前回使った乱数の a 倍に c を加えたものを M で割った余りを使えばよい。

例 11.1 $\langle M=16, a=5, c=1 \text{ の線形合同法} \rangle$ 初期値を $X_0=0$ とすると、 $X_1=5 \times 0+1=1, X_2=5 \times 1+1=6, X_3=5 \times 6+1=31=15 \pmod{16}$, となる。 X_4 以降も同様にして計算すると、次のように 12, 13, 2, 11, 8, 9, 14, 7, 4, 5, 10, 3, が得られ、 $X_{16}=0$ となつてもともどり、これから後は同じ数列(0, 1, 6, 15, …, 10, 3)の繰り返しになる。

$$\begin{array}{ll} X_1=5 \times 0+1=1 \pmod{16} & X_9=5 \times 8+1=9 \pmod{16} \\ X_2=5 \times 1+1=6 \pmod{16} & X_{10}=5 \times 9+1=14 \pmod{16} \\ X_3=5 \times 6+1=15 \pmod{16} & X_{11}=5 \times 14+1=7 \pmod{16} \\ X_4=5 \times 15+1=12 \pmod{16} & X_{12}=5 \times 7+1=4 \pmod{16} \\ X_5=5 \times 12+1=13 \pmod{16} & X_{13}=5 \times 4+1=5 \pmod{16} \\ X_6=5 \times 13+1=2 \pmod{16} & X_{14}=5 \times 5+1=10 \pmod{16} \\ X_7=5 \times 2+1=11 \pmod{16} & X_{15}=5 \times 10+1=3 \pmod{16} \\ X_8=5 \times 11+1=8 \pmod{16} & X_{16}=5 \times 3+1=0 \pmod{16} \end{array}$$

法 M の選択 発生する $\{X_n\}$ は 0 以上で $M-1$ 以下のいずれかの整数値をとることは明らかである。したがって 11.1 節で述べたような 10 進 1 けたの乱数列を作るためには $M=10$ とすればよいと思うかもしれないが、それではうまくいかない。それは、漸化式(11.1)によって生成される数列は周期的であり、その周期は M を超えられないからである。実際、それとわかる周期性は最もランダム性に反するものである。一般に、周期は実際に使う乱数の個数よりずっと長くないと困るので、 M はできるだけ大きく選ぶのがふつうである。

この場合、 $U_n=X_n/M$ は区間 $[0, 1)$ 上で連続型の一様分布

$$(11.2) \quad f(x)=1, \quad (0 \leq x < 1), \quad 0(\text{それ以外}) \quad (\text{密度関数})$$

$$F(x)=\begin{cases} 0 & (x < 0) \\ x & (0 \leq x < 1) \\ 1 & (x \geq 1) \end{cases} \quad (\text{累積分布関数})$$

*) $\text{mod } M$ は、英語では 'modulo(モジュロ) M ' と読む。

をする確率変数の実現値系列の近似値と見なせる。たとえば 10 進 1 けたの一様乱数が必要な場合には、 $10(X_n/M)$ の整数部分を使うことすればよい。

乗数 a の選択 法 M の値を定めたときに、(11.1)によって発生される数列 $\{X_n\}$ の周期が可能な最大値 M になるために乗数 a 、加数 c が満たすべき必要十分条件はよく知られていて、次の 3 条件がすべて成り立つことである。

- i) c が M と互いに素
- ii) $b=a-1$ が M を割り切るすべての素数の倍数
- iii) M が 4 の倍数であれば、 b も 4 の倍数

たとえば、例にのべた $M=16, a=5(b=4), c=1$ はこの条件を満たす。したがって、 $c=0$ とすると周期は最大値 M にはなりえないが、乱数の発生に要する時間を少し短くできるという理由でそうすることが多い。こうして得られる乱数発生法

$$(11.3) \quad X_n=aX_{n-1} \pmod{M}, \quad n \geq 1$$

は乗算型合同法とよばれ、その周期については次のことがわかっている。

i) M が素数の場合、可能な最長周期は $M-1$ であり、これを達成できるのは、 $X_0 \neq 0 \pmod{M}$ 、かつ $M-1$ の任意の素因数 p に対して

$$a^{(M-1)/p} \neq 1 \pmod{M}$$

が成り立つときに限られる。

1 語 = 32 ビットの計算機用としては、 $M=2^{31}-1$ という素数がよく使われていて、乗数 a としては $a=16807 (=7^5), 314159269, 630360016, 950706376$ などが使われる。ただし、IBM 360/370 系の計算機用のプログラムパッケージでよく使われている $a=16807$ は、ランダムネスに難点があるといわれている。

ii) $M=2^e (e \geq 4)$ を用いる場合は、可能な最長周期は $M/4$ であり、これを達成できるのは $a \pmod{8}=3$ または 5 で、かつ X_0 が奇数のときに限られる。

このとき、発生される乱数 X_n はすべて奇数であるが、すべての奇数が発生されるわけではないことに注意する必要がある。実際には、周期が $M/4$ のサイクルが 2 つあり、 X_0 の選び方によりいずれか一方のサイクルが実現する。

具体的には、たとえば章末の表 11.5 に示すパラメータの値を使えばよい。これらの数値が選ばれた根拠については、章末の付節を参照されたい。

※天文学のようなロマン※

大きい素数を探すという問題は、実は途方もなく大変である。実際、
 $2^{31}-1=2147483647$

が素数であることを証明するのは、簡単ではない。

一般に、 $M_p=2^p-1$ (p は正整数)の形をした素数は昔からメルセンヌ素数 Mersenne prime といわれている。まず、最初のいくつか

$$2^2-1=3, \quad 2^3-1=7, \quad 2^5-1=31, \quad 2^7-1=127$$

は素数である。だが、困ったことに

$$2^{11}-1=2047=23\cdot89$$

は素数でない。実際、 M_p が素数となるためには、 p 自身が素数であることが必要であるが、十分ではない(!)から、大変である。

M_p がメルセンヌ素数となるための p を探す長い——まるで天文学のような——歴史がある。まず、古代から18世紀までに

$$p=2, 3, 5, 7, 13, 17, 19, 31$$

が出そろう。ついで「ルーカスの素数判定法」の発見で

$$p=61, 89, 107, 127$$

がつけ加わった後、コンピュータで

$$p=521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213$$

まで発見された、と記載されている(岩波『数学辞典』第2版、1968)。だが、次々と作り出される高速コンピュータとその利用方法の進歩で、さらに

$$p=19937, 21701, 23209, 44497$$

がつけ加えられた(同、第3版、1985)。真ん中の2つは1978年に米国の高校生が見つけたものである。辞典の利用も、奥付け(本の最後にある著者、出版社、発行・印刷年月日などを記載した欄)まで調べないと、安心できないわけである。

これでもまだ終りではない。超高速コンピュータを長時間用いて $p=86243$ ('82年), 110503 ('88年), 216091 ('85年)が次々に発見され、さらに本書の校正中に、英国の科学者グループが $p=756839$ を見つけたとの報道があった(朝日新聞'92年3月28日夕刊)。この素数は227832桁で、「今日知られている最大のメルセンヌ素数」ということになる。(メルセンヌ素数以外なら、もっと巨大な 650801 桁の素数 $391581 \times 2^{2161893}-1$ が'89年に発見されている)。

ここまで来ると、何かロマンを感じてしまうが、どうだろうか。

例 11.2 <乗算型合同法> 乗算型合同法乱数列の最初の30項

$$X_i, \quad X_i/M \quad (1 \leq i \leq 30)$$

の計算結果は表 11.2, 11.2' のとおりである*). いずれの場合も $X_0=1$ であり、したがって X_1 が乗数 a を表す。また $M=2^{31}-1$ である。

表 11.2 乗算型合同法乱数列の例(例 11.2 を参照)

i	X_i	X_i/M	X_i	X_i/M
1	16807	0.7826369E-05	314159269	0.1462918E+00
2	282475249	0.1315378E+00	2132447592	0.9929983E+00
3	1622650073	0.7556053E+00	651647049	0.3034468E+00
4	984943658	0.4586501E+00	166574159	0.7756714E-01
5	1144108930	0.5327672E+00	360247326	0.1677532E+00
6	470211272	0.2189592E+00	551127943	0.2566389E+00
7	101027544	0.4704462E-01	327220579	0.1523740E+00
8	1457850878	0.6788647E+00	1438437321	0.6698246E+00
9	1458777923	0.6792964E+00	1750600443	0.8151869E+00
10	2007237709	0.9346929E+00	1827824785	0.8511472E+00
11	823564440	0.3835021E+00	1276617800	0.5944715E+00
12	1115438165	0.5194164E+00	525867184	0.2448760E+00
13	1784484492	0.8309653E+00	8832382	0.4112898E-02
14	74243042	0.3457211E-01	314941823	0.1466562E+00
15	114807987	0.5346163E-01	502164764	0.2338387E+00
16	1137522503	0.5297002E+00	189564492	0.8827285E-01
17	1441282327	0.6711494E+00	1262089744	0.5877063E+00
18	16531729	0.7698186E-02	1377296512	0.6413537E+00
19	823378840	0.3834156E+00	473834975	0.2206466E+00
20	143542612	0.6684224E-01	2120249050	0.9873179E+00
21	896544303	0.4174860E+00	1113864160	0.5186834E+00
22	1474833169	0.6867727E+00	1536715699	0.7155890E+00
23	1264817709	0.5889766E+00	1258932791	0.5862362E+00
24	1998097157	0.9304365E+00	588145047	0.2738764E+00
25	1817129560	0.8461669E+00	614242102	0.2860288E+00
26	1131570933	0.5269288E+00	1480833002	0.6895666E+00
27	197493099	0.9196489E-01	1528429405	0.7117304E+00
28	1404280278	0.6539190E+00	1377289104	0.6413502E+00
29	893351816	0.4159994E+00	1054243571	0.4909204E+00
30	1505795335	0.7011906E+00	1838194846	0.8559762E+00

$X_0=1, M=2^{31}-1$ としてある。なお、 a は X_1 を見よ。

*) $E+02, E-01$ などは $10^2, 10^{-1}$ を表す(浮動小数点形式)。

表11.2' 続き

<i>i</i>	X_i	X_i/M	X_i	X_i/M
1	630360016	0.2935342E+00	950706376	0.4427072E+00
2	1549035330	0.7213258E+00	129027171	0.6008296E-01
3	264620982	0.1232237E+00	1728259899	0.8047837E+00
4	529512731	0.2465736E+00	365181143	0.1700507E+00
5	1896697821	0.8832188E+00	1966843080	0.9158827E+00
6	2116530888	0.9855865E+00	1045174992	0.4866975E+00
7	1923129168	0.8955268E+00	636176783	0.2962429E+00
8	1674201058	0.7796106E+00	1602900997	0.7464089E+00
9	108088067	0.5033243E-01	640853092	0.2984205E+00
10	859154222	0.4000749E+00	429916489	0.2001955E+00
11	1946499387	0.9064094E+00	1671481929	0.7783445E+00
12	1377890442	0.6416302E+00	1285607481	0.5986576E+00
13	1382793310	0.6439133E+00	1066192246	0.4964845E+00
14	768302678	0.3577688E+00	48796904	0.2272283E-01
15	1014576563	0.4724490E+00	1176434418	0.5478200E+00
16	514017889	0.2393582E+00	776417870	0.3615477E+00
17	2050350098	0.9547687E+00	861463458	0.4011502E+00
18	1928578391	0.8980643E+00	1543924916	0.7189460E+00
19	863848128	0.4022606E+00	557508687	0.2596102E+00
20	246801402	0.1149259E+00	1650650964	0.7686443E+00
21	166165530	0.7737686E-01	741730640	0.3453952E+00
22	709020555	0.3301634E+00	1050856373	0.4893431E+00
23	1212767693	0.5647390E+00	678231403	0.3158261E+00
24	557139727	0.2594384E+00	1987866110	0.9256723E+00
25	1326452667	0.6176776E+00	622137510	0.2897054E+00
26	1278446513	0.5953230E+00	293824391	0.1368226E+00
27	1859909846	0.8660880E+00	516167496	0.2403592E+00
28	193562198	0.9013442E-01	1529740675	0.7123410E+00
29	1946383823	0.9063556E+00	599496926	0.2791625E+00
30	1393274952	0.6487942E+00	1297605210	0.6042445E+00

11.2.2 M系列に基づく方法

ここまで説明した線形合同法は簡便な方法である。ただし

- i) 使用する計算機の語長(16ビット, 32ビット等)によって周期が制約されるので、大規模なシミュレーションには向かない。
- ii) 多次元疎結合構造(章末付節を見よ)があるため、多次元分布の精度が制約されるという欠点がある。そこで、これらの欠点のない乱数発生法がいろいろ研究されているが、そのうちで性質が比較的よく解明されているものを紹介する。

合同法乱数列が上記の(ii)の欠点を持っているのは、章末付節に説明してある。

11.2 一様乱数の発生法

るとおり、使用している漸化式(11.1)の次数が1である(言いかえると、今回使う乱数 X_n が前回使った乱数 X_{n-1} から完全に決まってしまう)ためである。したがって、この欠点のない乱数列を発生するためには、次数が2以上の漸化式を使う必要がある。ここで紹介する方法は、次の漸化式を使うものである。

$$(11.4) \quad X_n = X_{n-q} \oplus X_{n-p} \quad (q < p)$$

ここで \oplus という記号は2進法で表した整数のけた上りなしの足し算

$$0 \oplus 0 = 0, \quad 0 \oplus 1 = 1, \quad 1 \oplus 0 = 1, \quad 1 \oplus 1 = 0$$

を意味し、たいていの計算機(あるいはプログラム言語)に備わっている排他的論理和exclusive orをとる演算によって実現できる(Pascalではxor, Cでは[^]などと書き表される)。パラメータ p, q の値は勝手にきめてよいわけではなく、表11.3に示すような値を選ぶのがふつうである。

A または B というとき、 A, B がともに成立することも含む場合を'inclusive or', 含まずこれを除外する場合を'exclusive or'といいう。

漸化式(11.4)の次数は p であるから、これに基づく乱数発生を計算機で実行するためには、長さ p の配列arrayを用意し、ここにこの乱数列のあい続く p 個の要素を貯えておく必要がある。この配列に最初に書き込む初期値 X_0, X_1, \dots, X_{p-1} は、原理的には次のように定める。

(11.4)式と同じ p, q を使って、漸化式

$$(11.5) \quad a_n = a_{n-q} + a_{n-p} \quad (\text{mod } 2)$$

によって1ビットの(0または1からなる)数列を作る。この数列は、M系列と呼ばれている。ただし初期値 a_0, a_1, \dots, a_{p-1} はランダムに(ただし、「すべてが0」とはならないように)定める。具体的には、たとえば合同法乱数を使って、各 a_i が $1/2$ ずつの確率で0か1をとるようにすればよい。ここで、話を簡単にするために、発生させる乱数 X_n のビット長 l は2の整数べき乗(16とか32

表11.3 M系列乱数の発生に使われる p, q の例

p	q
127	1, 7, 15, 30, 63
521	32, 48, 158, 168
607	105, 147, 273
1279	216, 418

乱数発生の漸化式に用いる p, q の値は勝手に決めるのではなく、この表を用いる。

など)であると仮定する。

まず X_0 は、M系列の最初の l 個の要素 a_0, a_1, \dots, a_{l-1} をこの順に並べて得られる2進整数とする。次の l 個の要素 $a_l, a_{l+1}, \dots, a_{2l-1}$ を並べて X_1 とし、以下同様にして X_{p-1} まで定める。なお、使用する計算機の語長は16あるいは32ビットであるが、負の乱数は発生させたくないで先頭の1ビットは使用しないことに対する場合には、上記の設定方法において先頭のビットを0で置き換えるべき。また、上記の初期値の設定は論理演算を用いて高速に行うことが可能である^{*}。

このようにして初期値を設定して得られる乱数列には、合同法乱数列のような多次元疎結晶構造ではなく、 l ビットの精度で見て p/l 以下の任意の次元で一様に分布することが知られている。

例11.3 <M系列による乱数発生例> 簡単な例で、以上の手続きを実行してみよう。 $p=5, q=3$ として、 X_n は2ビットの整数とする。漸化式(11.5)を使うための初期値は、表11.1の乱数表の一番上の行を使って定めよう。すなわち、10進2桁の整数 82, 69, 41, 01, 98 が並んでいると考え、それが50未満なら $a_n=0$ 、50以上なら $a_n=1$ とする。したがって、 $a_0=1, a_1=1, a_2=0, a_3=0, a_4=1$ となる。 a_5 以降は、漸化式 $a_n = a_{n-5} + a_{n-3} \pmod{2}$ を使って計算する。つぎに、 a_0, a_1 を並べて2ビットの整数 $X_0=11$ を作る。これは10進では3である。同様に、 $X_1=a_2a_3=00$ で、これは10進で0となる。以下同様にして、表11.4の X_4 までが得られる。 X_5 以降は、漸化式(11.4) $X_n = X_{n-5} \oplus X_{n-3}$ を使って作る。たとえば

$$\begin{array}{lll} X_0=11 & X_1=00 & X_2=11 \\ X_2=11(\oplus) & X_3=11(\oplus) & X_4=10(\oplus) \\ X_5=00 & X_6=11 & X_7=01 \end{array}$$

である。 X_{31} 以降を計算してみると、それは X_0 からの繰り返しと同じであり、この数列の周期は $31 (=2^p-1)$ であることがわかる。

なお、1周期のあいだに現れる数値(10進)の度数分布を調べてみると、0が7回、1, 2, 3 がそれぞれ8回で、ほぼ一様分布をしている。また、ひき続く2個の数値を座標成分とする点(X_n, X_{n+1})、 $0 \leq n \leq 30$ 、を方眼紙の上にプロット

^{*}) その方法に興味がある場合には、巻末伏見(32-35頁)を参照。

してみれば、これは格子点上にほぼ一様分布をしていることが確かめられるであろう。これは、線形合同法による乱数列の2次元分布(図11.3, 333頁)と著しく異なるところである。

11.3 正規乱数の発生法

一様分布以外の分布に従う確率変数の実現値(乱数)は、区間(0, 1)上の一様乱数に適当な変換を施して作るのがふつうである。正規分布は統計学では重要な確率分布なので、正規分布に従う正規乱数の発生のために、きわめてたくさんの変換法が提案されている。ここでは比較的単純な変換法を説明する^{*}。

なお、一般の正規分布 $N(\mu, \sigma^2)$ に従う乱数 Z は、標準正規分布 $N(0, 1)$ に従う乱数 X があれば、線形変換 $Z = \mu + \sigma X$ によって作ることができるので、ここでは X の発生のさせ方だけを考えればよい。以下では、 U_1, U_2 あるいは U 等は区間(0, 1)上の一様分布に従う互いに独立な乱数(一様乱数)を表す。

極座標法 一様乱数 U_1, U_2 から変換

$$(11.6) \quad X_1 = \sqrt{-2 \log U_1} \cdot \cos(2\pi U_2), \\ X_2 = \sqrt{-2 \log U_1} \cdot \sin(2\pi U_2)$$

によって得られる X_1 と X_2 は、互いに独立に標準正規分布 $N(0, 1)$ に従う。ここに \log は自然対数 \log_e を表す。これを極

表11.4 M系列乱数の例

$p=5, q=3$ の場合である。周期は31、ひきづく2個の数値を座標成分としてプロットすると、格子点上にほぼ一様分布をする。

n	X_n	
	2進	10進
0	11	3
1	00	0
2	11	3
3	11	3
4	10	2
5	00	0
6	11	3
7	01	1
8	11	3
9	01	1
10	01	1
11	00	0
12	00	0
13	10	2
14	01	1
15	01	1
16	10	2
17	01	1
18	11	3
19	11	3
20	00	0
21	01	1
22	10	2
23	11	3
24	10	2
25	10	2
26	10	2
27	00	0
28	01	1
29	00	0
30	10	2

^{*}) 非常に多数の乱数を使用する実験をする必要があり、乱数の発生に要する時間を少しでも短縮したい場合には、巻末の参考文献[1], [2]などを参照。

座標法あるいはボックス・ミュラー法という。

この方法が正しいことは、次のようにして確かめられる。いま、 X, Y を互いに独立に標準正規分布に従う確率変数とすると、これらの同時確率分布 $f(x, y)$ の確率エレメント(微小確率) $P(x \leq X \leq x+dx, y \leq Y \leq y+dy)$ は、 dx, dy を微小量として

$$(11.7) \quad f(x, y) dx dy = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dx dy = \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx dy$$

である。なお、 $dw = dx dy$ は直交座標における面積エレメント(微小面積)である。ここで、極座標を用いて

$$x = r \cos \theta, \quad y = r \sin \theta$$

とすると、極座標での面積エレメントは $dw = r dr \cdot d\theta$ であるから、(11.7) に対応する確率エレメントは

$$(11.8) \quad \frac{1}{2\pi} e^{-r^2/2} r dr \cdot d\theta$$

となる。ここで

$$s = r^2, \quad ds = 2r dr$$

とおいてみると、(11.8) は

$$(11.9) \quad \frac{1}{2\pi} e^{-s/2} \frac{ds}{2} d\theta = \frac{1}{2\pi} d\theta \cdot \frac{1}{2} e^{-s/2}$$

となる。 $1/(2\pi)$ は区間 $(0, 2\pi)$ 上の連続一様分布の密度関数、 $(1/2)e^{-s/2}$ は期待値 2 の指数分布 $Ex(1/2)$ の密度関数である。

したがって、 θ および R^2 をこれらの分布に従う互いに独立な確率変数とすると、

$$X = R \cos \theta, \quad Y = R \sin \theta$$

は互いに独立に標準正規分布をすることになる。これから(11.6) が出る。実際、区間 $(0, 1)$ 上の一様乱数から θ を得るのには $\theta = 2\pi U_2$ とすればよく、また、 $-2 \log U_1$ が指数分布 $Ex(1/2)$ に従うこととは比較的よく知られている(11.4 節参照)。

中心極限定理の応用 互いに独立な n 個の一様乱数 U_1, U_2, \dots, U_n の和を標準化して得られる。

$$(11.10) \quad X = \frac{U_1 + U_2 + \dots + U_n - (n/2)}{\sqrt{n/12}}$$

とすれば、 X の確率分布が $n \rightarrow \infty$ のとき標準正規分布に近づくという事実を利用する。よく使われる $n=12$ である。他の方法に比べるとかなり遅いので、実用的ではない。

この定理は中心極限定理を実験によって“実感する”のに使うとよいであろう。(本シリーズ I『統計学入門』第 8 章参照)。

2 次元正規分布の場合 平均が μ_1, μ_2 、分散が σ_1^2, σ_2^2 、相関係数が ρ の 2 次元正規分布 $N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2, \rho)$ に従う乱数 Y_1, Y_2 は、標準正規分布 $N(0, 1)$ に従う互いに独立な乱数 X_1, X_2 から

$$(11.11) \quad Y_1 = \mu_1 + \sigma_1 X_1, \quad Y_2 = \mu_2 + \sigma_2 (\rho X_1 + \sqrt{1-\rho^2} X_2)$$

のようにして生成される。この方法の正当性を確かめることは読者にまかせよう。正規乱数の加重和は正規乱数であるという事実を使い、 Y_1, Y_2 の平均、分散、および共分散を求めればよい(前掲第 7 章)。

3 次元以上の正規分布に従う正規乱数の発生の方法^{*}もある。

11.4 一般の乱数の発生法

正規分布以外の分布に従う乱数の発生法についても非常に多数の提案がされている。発生させたい乱数を X 、その分布関数を $F(x)$ 、それが連続分布ならば密度関数を $f(x)$ で表すことにする。

11.4.1 逆関数法

乱数の変換のための一般的手法はいろいろあるが、考え方として単純なのがこの方法である。 $F(x)$ の逆関数

$$(11.12) \quad F^{-1}(y) = \inf \{x : F(x) \geq y\},$$

$$0 \leq y \leq 1$$

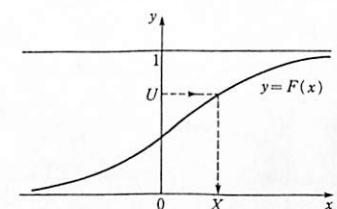


図 11.1 逆関数法

*) 卷末伏見(p. 70)参照。

を用いて

$$(11.13) \quad X = F^{-1}(U)$$

とすればよい(図11.1を参照せよ). X が $F(x)$ に従うことは、 $F^{-1}(y) \leq x$ と $y \leq F(x)$ が同値であることから、

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

でわかる。ここで、 $P(U \leq u) = u$ ($0 \leq u \leq 1$) であることを用いた。

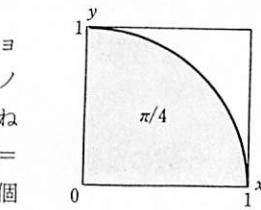
実際には、逆関数が解析的には求められなくて近似式を使う必要がある場合

※プリンセス・グレースのモンテ・カルロから※

フランスもイタリア国境近く、地中海に臨むリビエラ海岸にある小さな立憲君主国がモナコである。面積は 1.50 km^2 、人口は約3万人という極小の国家。この中心都市(首都ではない)がカジノと豪華ホテルの町、避寒地モンテ・カルロ Monte Carlo である。1956年4月、アメリカの映画女優グレース・ケリーが、公國の大公レーニエ3世と結婚して、現代版「シンデレラ結婚」と話題をもいた。若い人には、これよりも自動車レースで知られる。

「モンテ・カルロ法」は今では「乱数シミュレーション」とほとんど同じ意味で用いられている。カジノで行われるルーレット賭博に乱数のイメージを重ねたうまい命名法である。この方法は、正方形(1辺=1)中の4分円の面積 $\pi/4$ を、そこに落ちた乱数の個数から逆に求めることにより π を計算するなど、積分計算によく用いられる。(もっとも、 π を求める方法としては精度の良い方法ではない。)

だが、この格好のよい方法も、もともとは戦争に関係していた。歴史は第二次大戦中の1940年代にさかのぼる。原子物理学者フェルミ、ウラム、メトロポリス、フォン・ノイマン(後述)らは、連合国の大勝利のための原子爆弾開発プロジェクトにこの方法を精力的に用いた。原子物理学の中性子分布の計算にモンテ・カルロ法が必要だったのである。



モナコ
モンテ・カルロは地中海に臨む小公国モナコにある。この都市名に因んで、「モンテ・カルロ法」と名づけた。

も多く、また必ずしも高速な算法が得られるとは限らない。

例11.4 〈解析的な場合〉 逆関数 F^{-1} が解析的に求まるものとしては、指数分布: $F(x) = 1 - e^{-x}$, $x \geq 0$

ワイブル分布: $F(x) = 1 - \exp(-x^\alpha)$, $x \geq 0$

ロジスティック分布: $F(x) = \frac{1}{1 + e^{-x}}$, $-\infty < x < \infty$

コーリー分布: $F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} x$, $-\infty < x < +\infty$

などがある。たとえば、期待値1の指数分布 $Ex(1)$ の場合には、 $F^{-1}(y) = -\log(1-y)$ であるから

$$X = -\log(1-U)$$

とすればよい。なお、 $1-U$ は U と同じく区間(0,1)で一様分布に従うから、

$$(11.14) \quad X = -\log U$$

でもよい(ただし、 $U=0$ となる可能性がある場合には注意する必要がある)。期待値が μ の指数分布に従う乱数が必要なときは、 $X = -\mu \log U$ とすればよい。実際、この方法は、11.3節で述べた正規乱数を発生させるための極座標法の(11.9)で用いている。以上のように、指数分布に従う乱数を指數乱数という。

標準正規分布の場合には、逆関数 $F^{-1}(y)$ が解析的には求まらないが、その近似式はいろいろ提案されている。一例として

$$(11.15) \quad F^{-1}(y) = \begin{cases} -w, & 0 < y < 0.5 \text{ のとき} \\ +w, & 0.5 \leq y < 1 \text{ のとき} \end{cases}$$

ただし、

$$(11.16) \quad w = \left\{ z \left(2.0611786 - \frac{5.7262204}{z+11.640595} \right) \right\}^{1/2},$$

$$(11.17) \quad z = -\log[4y(1-y)]$$

がある(山内による近似式)。この近似式の相対誤差は次のようになる。

$$\left| \frac{\text{(求めた } F^{-1}(y) \text{)} - \text{(正しい } F^{-1}(y) \text{)}}{\text{正しい } F^{-1}(y)} \right| \leq 4.9 \times 10^{-4}$$

逆関数法は離散分布にも適用できる。一例として、幾何分布

$$(11.18) \quad P(X=n) = p(1-p)^{n-1}, \quad n=1, 2, \dots$$

をとりあげてみよう。この分布の分布関数は

$$(11.19) \quad F(n) = 1 - \sum_{k=n+1}^{\infty} p(1-p)^{k-1} = 1 - (1-p)^n$$

であるから、

$$(11.20) \quad (1-p)^{n-1} \geq U > (1-p)^n$$

を満たす n を X の値とすればよい。すなわち

$$(11.21) \quad X = \left\lceil \frac{\log U}{\log(1-p)} \right\rceil + 1$$

とすればよい。ただし、 $\lfloor v \rfloor$ は v の整数部分を表す。

幾何分布は逆関数法によって比較的効率のよい乱数発生ができる例であるが、その他の離散分布の場合には、分布関数あるいはその逆関数が簡単な式で表せない、効率のよい発生法が得られないことが多い。

つぎに示す別名法は、このような場合に使うと便利なことが多い。

11.4.2 別名法

取りうる値の個数が有限個の任意の離散分布

$$(11.22) \quad P(X=x_k) = p_k \quad (1 \leq k \leq K)$$

に従う乱数を発生させるためにウォーカー(Walker, A. J.)が考案した巧妙な方法である。逆関数法と違って、 K がいくら大きくても 1 個の乱数の発生に必要な比較演算が 1 回だけであるという特長を持つ。ただし、そのために最初に次のようにして表を用意しておく必要がある。

[初期設定] 次の手順に従って $a_k, v_k (1 \leq k \leq K)$ を求める。

$$1^\circ \quad v_k \leftarrow Kp_k \quad (1 \leq k \leq K)$$

2° $v_k \geq 1$ を満たす k の集合を G , $v_k < 1$ を満たす k の集合を S とする。

3° S が空集合でない限り、3.1° から 3.4° までを繰り返し実行する。

3.1° S, G から 1 つずつ任意の要素を選ぶ。それを k, l とする。

$$3.2^\circ \quad a_k \leftarrow x_l, \quad v_l \leftarrow v_l - (1 - v_k)$$

3.3° S から k を除く。

3.4° $v_l < 1$ なら l を G から S に移す。

以上で準備が終る。乱数が必要になるたびに次の手順に従い 1 個を発生する。

【乱数発生】

1° $(0, K)$ 上の一様乱数 $V = KU$ を発生する。

2° $k \leftarrow \lfloor V \rfloor + 1, u \leftarrow k - V$

3° $u \leq v_k$ ならば $X \leftarrow x_k$, そうでなければ $X \leftarrow a_k$

この方法の基礎になっているのは、次の事実である：

“任意の離散分布(11.22)は、 K 個の 2 点分布の等加重和として表現できる。”(ただし、表現のしかたは一意的ではない。)

上の算法の記号を用いて言えば、 k 番目の 2 点分布は、確率 v_k で値 x_k を取り、確率 $1 - v_k$ で値 x_k (a_k の別名 alias) をとることになる。[乱数発生] の手順 1°, 2° では K 個の 2 点分布のうちのいずれか 1 つ(k)が等確率で選ばれる。また、[初期設定] の手順 3° では、確率の小さい x_k に対して確率の大きい x_l を a_k として組み合わせて 2 点分布を構成している。

例 11.5 〈別名法の原理〉 簡単な例を示そう。 $K=5$ とし、 $p_k (1 \leq k \leq 5)$ は 0.12, 0.24, 0.38, 0.16, 0.10 としよう。 Kp_k を柱状グラフで表せば、図 11.2(a) のようになる。 Kp_k の和は K であるから、柱状グラフの高さの平均値は 1 である。そこで、1 より高いところを削って 1 より低いところを埋め、すべての柱の高さを 1 にそろえることにしよう。もちろん、そのやり方は、何通りもあるが、たとえば図 11.2(b) のようにすればよい。そうすると、もとの離散分

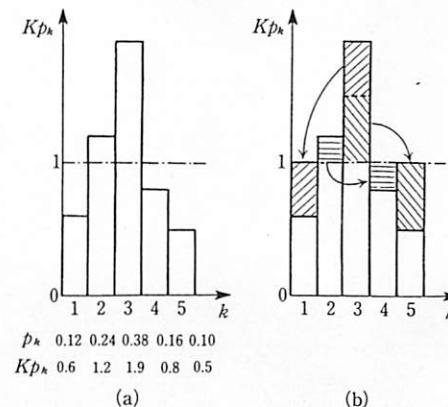


図 11.2 別名法

布は、次の5つ($k=1, 2, 3, 4, 5$)の2点分布の等加重(=1/5)として表現できたことになる。ただし、*のところは値が定まらないが、これをとる確率は0なのでさしつかえない。

11.4.3 ガンマ分布

ガンマ分布に従う乱数をガンマ乱数という。ガンマ分布 $Ga(\alpha, 1/\lambda)$ は

$$(11.23) \quad f(x) = \frac{1}{\lambda \Gamma(\alpha)} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-x/\lambda}, \quad x \geq 0$$

なる密度関数で与えられる。期待値は $\alpha\lambda$ 、分散は $\alpha\lambda^2$ となる。(11.23)におけるこの α, λ を形状パラメータ、尺度パラメータという(統計学に用いるときは、「パラメータ」を「母数」という)。このとき、3つの重要な性質：

- i) ガンマ分布 $Ga(1, 1/\lambda)$ は期待値が λ の指数分布 $Ex(1/\lambda)$,
- ii) 標準正規分布 $N(0, 1)$ に従う確率変数 Z の平方 Z^2 は、ガンマ分布 $Ga(1/2, 1/2)$ に従う,
- iii) X_1, X_2 が互いに独立で、共通の尺度パラメータの値 λ をもつガンマ分布 $Ga(\alpha_1, 1/\lambda), Ga(\alpha_2, 1/\lambda)$ にそれぞれ従うならば、 $X_1 + X_2$ も再び同一の尺度パラメータの値 λ をもつガンマ分布 $Ga(\alpha_1 + \alpha_2, 1/\lambda)$ に従う(ガンマ分布の再生性)

がある^{*)}。iii)は一般的なものであり、i), ii)は $Ga(1, 1/\lambda), Ga(1/2, 1/2)$ についてのものである。したがって、

a) 形状パラメータが整数 $\alpha = n$ のときは、 n 個の独立な一様乱数から

$$(11.24) \quad X = -\lambda \log(U_1 U_2 \cdots U_n)$$

とすれば、(11.14), i), iii)から、 X は $Ga(n, 1/\lambda)$ に従うガンマ乱数を与える。

b) 形状パラメータが半整数 $\alpha = n + (1/2)$ となっているときは、 Z^2 が $Ga(1/2, 1/2)$ に、 $(\lambda/2)Z^2$ が $Ga(1/2, 1/\lambda)$ に従うから、iii)から

$$(11.25) \quad X = -\lambda \log(U_1 U_2 \cdots U_n) + (\lambda/2)Z^2$$

が $Ga(n+1/2, 1/\lambda)$ に従うガンマ乱数を与える。

なお、自由度 k の χ^2 分布 $\chi^2(k)$ は、ガンマ分布 $Ga(k/2, 1/2)$ である。これは、 χ^2 分布の定義と、ii), iii)からただちに従うが、 k の偶奇 $k=2n, k=2n+1$

*) 本シリーズI『統計学入門』第6章および練習問題。

に応じて、a), b)に帰するから、 $\chi^2(k)$ に従う χ^2 乱数も発生できる。

α が整数または半整数でない場合、あるいは α が大きい場合には、この方法は効率的でないので、代わりの方法がいろいろ提案されている。

X がガンマ分布 $Ga(\alpha, 1)$ に従って分布するとき、 α が大きければ、

$$(11.26) \quad Y = \frac{\left(\frac{X}{\alpha}\right)^{1/3} - \left(1 - \frac{1}{9\alpha}\right)}{\sqrt{1/9\alpha}}$$

の確率分布は標準正規分布に近いことが知られている。これを X について解くと

$$(11.27) \quad X = \alpha \left(\frac{Y}{\sqrt{9\alpha}} + 1 - \frac{1}{9\alpha} \right)^3$$

となる。したがって、標準正規乱数 Y に変換(11.27)を施せば、 X は近似的に $Ga(\alpha, 1)$ に従う乱数となる。 $Ga(\alpha, 1/\lambda)$ に従う乱数が必要な場合には、(11.27)の右辺を λ 倍すればよい。

以上の近似を、「ウィルソン・ヒルファーティの近似」という。

11.4.4 t 分布

Z が標準正規分布に従い、 Y が自由度 n の χ^2 分布 $\chi^2(n)$ に従い、 Z と Y が互いに独立ならば、

$$(11.28) \quad X = \frac{Z}{\sqrt{Y/n}}$$

は自由度が n の t 分布 $t(n)$ に従う。正規乱数および χ^2 乱数の発生方法はすでに述べたので、(11.28)により t 分布に従う乱数が発生できる。

11.4.5 ポアソン分布

平均が λ のポアソン分布

$$(11.29) \quad P(X=n) = e^{-\lambda} \frac{\lambda^n}{n!}, \quad n=0, 1, 2, \dots$$

に従う乱数を発生する方法のひとつは、ポアソン分布と指数分布の間に成り立つ次の関係を利用するものである：

“ある現象の生起間隔が期待値 $1/\lambda$ の指数分布 $Ex(\lambda)$ に従うならば、この現象が単位時間内に起こる回数は期待値 λ のポアソン分布 $Po(\lambda)$ に従う。”

したがって、 X_1, X_2, \dots を互いに独立で期待値 $1/\lambda$ の指数乱数とし、不等式

$$(11.30) \quad X_1 + X_2 + \dots + X_n < 1$$

を満たす最大の n を X の取る値とすればよい。指数乱数は逆関数法で作ることにすると、条件(11.30)は

$$(11.31) \quad U_1 U_2 \cdots U_n > e^{-\lambda}$$

と同等である。

λ が大きいときには、上の方法は効率的でないので、次のような近似を用いるのがよい。まず第一は、中心極限定理による近似である。期待値 λ のポアソン分布 $Po(\lambda)$ の分散は λ であるから、 λ が大きいときには $(X - \lambda)/\sqrt{\lambda}$ の分布は標準正規分布に近い。したがって、 Z を標準正規乱数として

$$(11.32) \quad X = \max \{0, \lfloor \lambda + \sqrt{\lambda} Z - 0.5 \rfloor\}$$

とすればよい。ここで $\lfloor v \rfloor$ は v の整数部分を表し、また 0.5 は離散分布を連続分布で近似するときの精度を上げるために用いられる不連続補正 correction for discontinuity とよばれる量である。

もうひとつの近似の方法は、分布を途中で打ち切って別名法を用いるものである。すなわち、 $P(X > n_0)$ が十分小さくなるような n_0 (たとえば $\lambda + 6\sqrt{\lambda}$ 程度の整数) を選び、 $X=0, 1, \dots, n_0$ に対する表をあらかじめ作成しておき、表を引くことによって乱数を発生する。 λ をたびたび変える必要がなければ、非常に効率のよい方法である。

付節 多次元疎結晶構造とスペクトル検定

線形合同法によって発生される乱数列の 1 周期分の中には同一の数はない。したがって、この数列の連続する 2 個の要素を座標成分とする点列

$$\{(X_n, X_{n+1}); n=0, 1, \dots\}$$

の 1 周期分をプロットすると、たとえば図 11.3 のようになり、座標軸に平行な任意の直線上にはたかだか 1 点しか載らないし、しかもすべての点は大変規則的に並んでしまう。もっと高い次元の空間に点をプロットする場合でも事情は同様で、 k 次元の場合には、 M^k 個の格子点のうちたかだか M 点にしかプロットされず、それらは $(k!M)^{1/k}$ 枚以下の等間隔に並んだ平行な $(k-1)$ 次元超平面上に規則的に並んでしまう。この超平面の枚数の上界は k が増すとともに小さくなり、たとえば $M=2^{32}$ の場

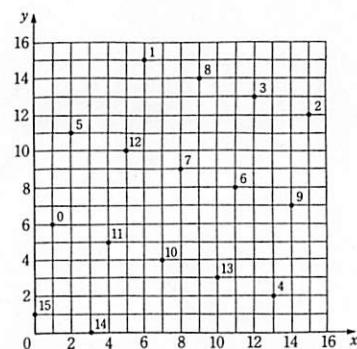


図 11.3 合同法乱数列の 2 次元疎結晶構造： $M=16$,

$$a=5, c=1, X_0=1$$

の場合。 $M=16, a=5, c=1, X_0=1$ とすると、(11.1)式によって発生される数列の 1 周期分は、 $1, 6, 15, \dots, 3, 0$ となり、 $15 (=M-1)$ 以下のすべての非負整数がちょうど 1 回ずつ現れる。この数列のあいにく 2 個の要素を座標成分とする点列 $\{P_n(X_n, X_{n+1}); n=0, 1, 2, \dots\}$ をプロットした。

合には、 $k=3$ なら 2953 , $k=6$ で 120 , $k=10$ で 41 となる。このような性質は多次元疎結晶構造とよばれています。合同法乱数列のひとつ欠点とされています。

点列が超平面群の上に載ってしまうことは避けられないにしても、超平面の間隔が小さければ、実用上は大きな支障はないであろうと考えられる。そこで、超平面の間隔を計算する算法が考案されていて、これによって乗数 a の良さを判定することをスペクトル検定とよんでいる ($c \neq 0$ の場合、 c の値は超平面の間隔には影響しない)。具体的には、 k 次元空間の超平面の間隔を求めるためには、ます

$$(11.33) \quad \nu_k = \min \left\{ \sqrt{s_1^2 + \dots + s_k^2}; \quad s_1 + as_2 + \dots + a^{k-1}s_k \equiv 0 \pmod{M'}, \right. \\ \left. (s_1, \dots, s_k) \text{ は非零整数ベクトル} \right\}$$

の値を計算する。ここで M' は、数列の周期が M または $M-1$ の場合は $M'=M$, $a \pmod{8}=5$ で周期が $M/4$ の場合には $M'=M/4$ とする。いずれの場合にも超平面の

表 11.5 スペクトル検定結果の例

a	c	M	$\log_2 \nu_2$	$\log_2 \nu_3$	$\log_2 \nu_4$	$\log_2 \nu_5$	$\log_2 \nu_6$
1664525	0	2^{32}	14.1	9.3	6.0	5.7	4.4
1664525	*	2^{32}	16.1	10.6	8.0	6.0	5.0
1566083941	0	2^{32}	14.8	9.7	7.5	5.6	4.2
1566083941	*	2^{32}	16.1	10.5	7.7	6.1	4.7
16807	0	$2^{31}-1$	14.0	9.3	7.2	6.1	4.9
314159269	0	$2^{31}-1$	15.2	9.9	7.6	5.9	5.1
630360016	0	$2^{31}-1$	15.3	9.3	7.6	6.2	4.7
950706376	0	$2^{31}-1$	15.4	10.3	7.8	6.2	5.3

c の列の * は、任意の奇数を表す。

※フォン・ノイマンは人間だったか？※

「モンテ・カルロ法」と命名したのは、今世紀きっての天才、近代計算機の発明者でもあったフォン・ノイマン J. von Neumann である。『ラス・ヴェガス法』などとしなかったのは、フェルミの故国イタリアに敬意を表したものか、定かではない。Monte Carlo とは、イタリア語で「カルロ(の)山」の意味である。1956年には、フロリダ大学で大々的にこの「モンテ・カルロ法」の会議(シンポジウム)が開催されるなど、今では数理全般で役に立つ一般的方法の名として広く知られている。

乱数を使う数値実験のアイデア自体はもっと古い。19世紀の「ピュッフォンの針」、20世紀に入ってからも、 t 分布の導出のために W. ゴセット(1908)が実験で用いるなど、アイデアが統計学者を引きつけてきた歴史は意外と長い。

フォン・ノイマンの名は、『量子力学の数学的基礎』など、物理学を志した人は一種のあこがれをもった経験があるだろうが、亡くなったとき「彼は‘人間’だったのだろうか」といわれたくらいに、かかわった広い分野のすべてで超一流の業績をあげている。

シャーロック・ホームズにも関心があったようだ。ホームズが、天才的老数学学者で希代の犯罪の大黒幕モリアティー教授と息詰まるような知的逃走劇^{*)}を演じる事件も、あまりにも有名なゲーム理論の原典『ゲーム理論と経済行動』Theory of Games and Economic Behavior(8章)で、くわしく解かれ分析されてしまった。「天才的老数学学者」という所に自ら異常な関心を持ったのかもしれないが、凄い人だったようである。

*) 『最後の問題』The Final Problem.



フォン・ノイマン
(1903-57)

間隔は M/ν_k である^{*)}。

ν_k は合同法乱数列の ‘ k 次元精度’ とよばれることがある。それは、この数列によって生成される k 次元点列は、その座標成分を上位 $\log_2 \nu_k$ ピットまでの分解能で見れば、ほぼ一様に分布しているものと見なせるからである。このピット数と乱数列の周期 N との間には次の関係があることが知られている。

(11.34)

$$\log_2 \nu_k \leq (\log_2 N)/k + c_k$$

*) 巻末クヌース参照。

ここに

$$c_2 = 0.104, \quad c_3 = 0.167, \quad c_4 = 0.25, \quad c_5 = 0.3, \quad c_6 = 0.368$$

である。

したがって、たとえば 3 次元空間内に 10 ピット程度(10 進法で 3 けた程度)の精度で点を分布させようとするならば、 $\log_2 N$ はほぼ 30 程度以上にする必要がある、いいかえると、(11.1)の整数演算は 30 ピット程度以上のけた数で行う必要がある。16 ピットのパソコンの BASIC の乱数発生関数 RND の整数演算は 16~24 ピット程度のけた数で行っているものが多いようであり、けた数不足であるといえよう。

スペクトル検定については、多くの計算結果が報告されているが、そのうちで比較的成績の良いものをごくわずかだけ選んで表 11.5 に示す。

◆◆◆◆ 第11章 練習問題 ◆◆◆◆

11.1 <既存の一様乱数発生プログラムのチェック> 身近にある計算機で使えるプログラム言語に備わっている一様乱数発生法がどのようなものであるかを調べてみよ。もし、線形合同法であることだけはわかったけれど、乗数 a 、加数 c 、法 M が不明である場合には、実際に発生された乱数列からこれらをつきとめる方法を工夫せよ。

11.2 <乗算型合同法による一様乱数発生プログラムの作成> 適当なプログラム言語を使って、例 11.2 で示したパラメータを用いる乗算型合同法による乱数発生のプログラムを作成せよ。これを用いて発生させた乱数列が表 11.2 のようになることを確かめよ。

11.3 <M 系列による乱数発生プログラムの作成> 適当なプログラム言語を使って、M 系列に基づく一様乱数発生のプログラムを作成せよ。これを使って 10,000 個の乱数を発生し、これが一様分布からのランダム・サンプルと見なせるかどうか検定(適合度の検定)せよ。

11.4 <極座標法> 正規乱数を発生するための極座標法(11.6)が正しいことを、次の方針に従って確かめよ。

確率変数の組 (U_1, U_2) の密度関数を $f(u_1, u_2)$ 、 (X_1, X_2) の密度関数を $g(x_1, x_2)$ とすると、

$$g(x_1, x_2) = |J| \cdot f(u_1, u_2)$$

が成り立つ。ただし、 $|J|$ はヤコビ行列式(ヤコビアン)

$$J = \frac{\partial(u_1, u_2)}{\partial(x_1, x_2)} = \begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{vmatrix}$$

の絶対値である。

11.5 <中心極限定理の実験的検証> 上の練習問題 11.2 または 11.3 で作成した一様乱数発生プログラムを使い、(11.10)により X を 1,000 個発生させよ。これが標準正規分布からのランダム・サンプルと見なせるかどうかを検定せよ。

統計數值表

付表 1 正規分布表(上側確率)――――――――――	335
付表 2 χ^2 分布表(パーセント点)――――――――――	336
付表 3 t 分布表(パーセント点)――――――――――	337
付表 4 F 分布表(パーセント点)――――――――――	338

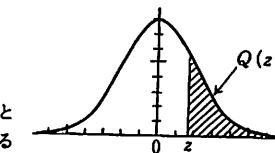
本書のために編者の一人藤野和建が計算、作成したものである。本表のほか、第I巻の統計数値表(日本規格協会「簡約統計数値表」による)を参照してもよい。その他の統計数値については、中山伊知郎『統計学大辞典』東洋経済新報社(1959)の巻末を見られたい。また、乱数表については、第11章表11.1を参照のこと。

エクセルによる統計数値表¹の作表も可能であるが、小数下位で差を生じることがある。

付表1 正規分布表(上側確率)

 $z=0 \sim 4.09$ に対する、標準正規分布の上側確率

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty \exp\left(-\frac{u^2}{2}\right) du$$

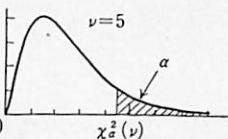
を与える。累積分布関数の値は $\phi(z) = 1 - Q(z)$ である。例. $z=1.96$ に対する上側確率 $Q(1.96)$ は、左の見出し 1.9 の行と上の見出し 0.06 の列の交点の値 0.02500 である。 $z=3.50$ に対する $Q(3.50)$ は $0.0^{233}=0.000233$ である。

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.0 ³ 968	.0 ³ 935	.0 ³ 904	.0 ³ 874	.0 ³ 845	.0 ³ 816	.0 ³ 789	.0 ³ 762	.0 ³ 736	.0 ³ 711
3.2	.0 ³ 687	.0 ³ 664	.0 ³ 641	.0 ³ 619	.0 ³ 598	.0 ³ 577	.0 ³ 557	.0 ³ 538	.0 ³ 519	.0 ³ 501
3.3	.0 ³ 483	.0 ³ 466	.0 ³ 450	.0 ³ 434	.0 ³ 419	.0 ³ 404	.0 ³ 390	.0 ³ 376	.0 ³ 362	.0 ³ 349
3.4	.0 ³ 337	.0 ³ 325	.0 ³ 313	.0 ³ 302	.0 ³ 291	.0 ³ 280	.0 ³ 270	.0 ³ 260	.0 ³ 251	.0 ³ 242
3.5	.0 ³ 233	.0 ³ 224	.0 ³ 216	.0 ³ 208	.0 ³ 200	.0 ³ 193	.0 ³ 185	.0 ³ 178	.0 ³ 172	.0 ³ 165
3.6	.0 ³ 159	.0 ³ 153	.0 ³ 147	.0 ³ 142	.0 ³ 136	.0 ³ 131	.0 ³ 126	.0 ³ 121	.0 ³ 117	.0 ³ 112
3.7	.0 ³ 108	.0 ³ 104	.0 ³ 996	.0 ³ 957	.0 ³ 920	.0 ³ 884	.0 ³ 850	.0 ³ 816	.0 ³ 784	.0 ³ 753
3.8	.0 ² 723	.0 ² 695	.0 ² 667	.0 ² 641	.0 ² 613	.0 ² 591	.0 ² 567	.0 ² 544	.0 ² 522	.0 ² 501
3.9	.0 ² 481	.0 ² 461	.0 ² 443	.0 ² 425	.0 ² 407	.0 ² 391	.0 ² 375	.0 ² 359	.0 ² 345	.0 ² 330
4.0	.0 ² 317	.0 ² 304	.0 ² 291	.0 ² 279	.0 ² 267	.0 ² 256	.0 ² 245	.0 ² 235	.0 ² 225	.0 ² 216

付表2 χ^2 分布表(パーセント点)

自由度 ν の χ^2 分布に対し、上側確率 α に対応する χ^2 の値を与える。この値を上側 100α パーセント点とよび、 $\chi_{\alpha}^2(\nu)$ で表す。

例. 自由度 $\nu=10$, $\alpha=0.05$ に対しては、 $\chi_{0.05}^2(10)=18.307$ である。

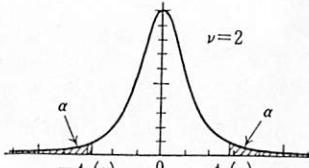


ν	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.039270	.0315709	.0298207	.0239321	.0157791	.45494	2.7055	3.8415	5.0239	6.6349	7.8794
2	.010025	.020101	.050636	.10259	.21072	1.3863	4.6052	5.9915	7.3778	9.2103	10.597
3	.071722	.11483	.21580	.35185	.58437	2.3660	6.2514	7.8147	9.3484	11.345	12.838
4	.20699	.29711	.48442	.71072	1.0636	3.3567	7.7794	9.4877	11.143	13.277	14.860
5	.41174	.55430	.83121	1.1455	1.6103	4.3515	9.2364	11.070	12.833	15.086	16.750
6	.67573	.87209	1.2373	1.6354	2.2041	5.3481	10.645	12.592	14.449	16.812	18.548
7	.98926	1.2390	1.6899	2.1673	2.8331	6.3458	12.017	14.067	16.013	18.475	20.278
8	1.3444	1.6465	2.1797	2.7326	3.4895	7.3441	13.362	15.507	17.535	20.090	21.955
9	1.7349	2.0879	2.7004	3.3251	4.1682	8.3428	14.684	16.919	19.023	21.666	23.589
10	2.1559	2.5582	3.2470	3.9403	4.8652	9.3418	15.987	18.307	20.483	23.209	25.188
11	2.6032	3.0535	3.8157	4.5748	5.5778	10.341	17.275	19.675	21.920	24.725	26.757
12	3.0738	3.5706	4.4038	5.2260	6.3038	11.340	18.549	21.026	23.337	26.217	28.300
13	3.5650	4.1069	5.0088	5.8919	7.0415	12.340	19.812	22.362	24.736	27.688	29.819
14	4.0747	4.6604	5.6287	6.5706	7.7895	13.339	21.064	23.685	26.119	29.141	31.319
15	4.6009	5.2293	6.2621	7.2609	8.5468	14.339	22.307	24.996	27.488	30.578	32.801
16	5.1422	5.8122	6.9077	7.9616	9.3122	15.338	23.542	26.296	28.845	32.000	34.267
17	5.6972	6.4078	7.5642	8.6718	10.085	16.338	24.769	27.587	30.191	33.409	35.718
18	6.2648	7.0149	8.2307	9.3905	10.865	17.338	25.989	28.869	31.526	34.805	37.156
19	6.8440	7.6327	8.9065	10.117	11.651	18.338	27.204	30.144	32.852	36.191	38.582
20	7.4338	8.2604	9.5908	10.851	12.443	19.337	28.412	31.410	34.170	37.566	39.997
21	8.0337	8.8972	10.283	11.591	13.240	20.337	29.615	32.671	35.479	38.932	41.401
22	8.6427	9.5425	10.982	12.338	14.041	21.337	30.813	33.924	36.781	40.289	42.796
23	9.2604	10.196	11.689	13.091	14.848	22.337	32.007	35.172	38.076	41.638	44.181
24	9.8862	10.856	12.401	13.848	15.659	23.337	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	24.337	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	25.336	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	26.336	36.741	40.113	43.195	46.963	49.645
28	12.461	13.365	15.308	16.928	18.939	27.336	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	28.336	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	29.336	40.256	43.773	46.979	50.892	53.672
31	14.458	15.655	17.539	19.281	21.434	30.336	41.422	44.985	48.232	52.191	55.003
32	15.134	16.362	18.291	20.072	22.271	31.336	42.585	46.194	49.480	53.486	56.328
33	15.815	17.074	19.047	20.867	23.110	32.336	43.745	47.400	50.725	54.776	57.648
34	16.501	17.789	19.806	21.664	23.952	33.336	44.903	48.602	51.966	56.061	58.964
35	17.192	18.509	20.569	22.463	24.797	34.336	46.059	49.802	53.203	57.342	60.275
36	17.887	19.233	21.336	23.269	25.643	35.336	47.212	50.998	54.437	58.619	61.581
37	18.586	19.960	22.106	24.075	26.492	36.336	48.363	52.192	55.668	59.893	62.883
38	19.289	20.691	22.878	24.884	27.343	37.335	49.513	53.384	56.896	61.162	64.181
39	19.996	21.426	23.654	25.695	28.196	38.335	50.660	54.572	58.120	62.428	65.476
40	20.707	22.164	24.433	26.509	29.051	39.335	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	49.335	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	59.335	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	69.334	85.527	90.531	95.023	100.43	104.21
80	51.172	53.540	57.153	60.391	64.278	79.334	96.578	101.88	106.63	112.33	116.32
90	59.196	61.754	65.647	69.126	73.291	89.334	107.57	113.15	118.14	124.12	128.30
100	67.328	70.065	74.222	77.929	82.358	99.334	118.50	124.34	129.56	135.81	140.17
110	75.550	78.458	82.867	86.792	91.471	109.33	129.39	135.48	140.92	147.41	151.95
120	83.852	86.923	91.573	95.705	100.62	119.33	140.23	146.57	152.21	158.95	163.65

付表3 t 分布表(パーセント点)

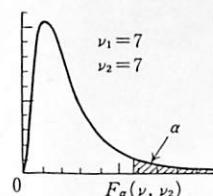
自由度 ν の t 分布に対し、上側確率 α (両側確率 2α)に対応する t の値を与える。この値を上側 100α パーセント点とよび、 $t_{\alpha}(\nu)$ で表す。

例. 自由度 $\nu=10$ の上側 5 パーセント点は $t_{0.05}(10)=1.812$ である。また、両側確率 5 パーセントの点は $t_{0.025}(10)=2.228$ である。



ν	α	.250	.200	.150	.100	.050	.025	.010	.005	.0010	.0005
	2α	.500	.400	.300	.200	.100	.050	.020	.010	.0020	.0010
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	318.309	636.619	
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599	
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924	
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610	
5	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869	
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959	
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408	
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041	
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781	
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587	
11	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437	
12	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318	
13	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221	
14	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140	
15	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073	
16	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015	
17	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965	
18	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922	
19	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883	
20	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850	
21	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819	
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792	
23	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768	
24	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745	
25	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725	
26	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707	
27	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690	
2											

付表4 F分布表(パーセント点)



自由度 v_1, v_2 の F 分布に対し、上側確率 α に対応する F の値を与える。この値を上側 100α パーセント点とよび、 $F_\alpha(v_1, v_2)$ で表す。

例1. 自由度 $v_1=4, v_2=10$ の上側 5 パーセント点は $F_{0.05}(4, 10)=3.478$ である。

自由度が表にないときの $F_\alpha(v_1, v_2)$ は、自由度の逆数による線形補間を用いるか、近似式により求める。線形補間による方法を例2, 3, 4 に、近似式を用いる方法を例5に示す。

$\alpha=0.05$

v_2	1	2	3	4	5	6	7	8	9	10	12
v_1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.413
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.745
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.912
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.678
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.000
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.575
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.284
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.073
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.913
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.788
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.687
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.604
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.534
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.475
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.425
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.381
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.342
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.308
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.278
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321	2.250
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.226
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275	2.204
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255	2.183
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.165
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.148
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204	2.132
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.118
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177	2.104
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.092
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153	2.080
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142	2.070
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133	2.060
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123	2.050
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114	2.041
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209	2.153	2.106	2.033
37	4.105	3.252	2.859	2.626	2.470	2.356	2.270	2.201	2.145	2.098	2.025
38	4.098	3.245	2.852	2.619	2.463	2.349	2.262	2.194	2.138	2.091	2.017
39	4.091	3.238	2.845	2.612	2.456	2.342	2.255	2.187	2.131	2.084	2.010
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	2.003
45	4.057	3.204	2.812	2.579	2.422	2.308	2.221	2.152	2.096	2.049	1.974
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026	1.952
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.917
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951	1.875
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.834
∞	3.841	2.996	2.605	2.372	2.214	2.099	2.010	1.938	1.880	1.831	1.752

例2. 自由度 $v_1=4, v_2=90$ の上側 5 パーセント点は、 $v_2=90$ が表に含まれていないので、 $v_2=80$ に対する $F_{0.05}(4, 80)=2.486$ 、および $v_2=120$ に対する $F_{0.05}(4, 120)=2.447$ 、さらに、

$$\left(\frac{1}{90} - \frac{1}{120}\right) / \left(\frac{1}{80} - \frac{1}{120}\right) = 0.666$$

より、

$$F_{0.05}(4, 90)=2.447+(2.486-2.447) \times 0.666=2.473$$

とする。

	14	16	18	20	24	30	40	60	120	∞	v_1	v_2
245.36	246.46	247.32	248.01	249.05	250.10	251.14	252.20	253.25	254.31	1		
19.424	19.433	19.440	19.446	19.454	19.462	19.471	19.479	19.487	19.496	2		
8.715	8.692	8.675	8.660	8.639	8.617	8.594	8.572	8.549	8.526	3		
5.873	5.844	5.821	5.803	5.774	5.746	5.717	5.688	5.658	5.628	4		
4.636	4.604	4.579	4.558	4.527	4.496	4.464	4.431	4.398	4.365	5		
3.956	3.922	3.896	3.874	3.841	3.808	3.774	3.740	3.705	3.669	6		
3.529	3.494	3.467	3.445	3.410	3.376	3.340	3.304	3.267	3.230	7		
3.237	3.202	3.173	3.150	3.115	3.079	3.043	3.005	2.967	2.928	8		
3.025	2.989	2.960	2.936	2.900	2.864	2.826	2.787	2.748	2.707	9		
2.865	2.828	2.798	2.774	2.737	2.700	2.661	2.621	2.580	2.538	10		
2.739	2.701	2.671	2.646	2.609	2.570	2.531	2.490	2.448	2.404	11		
2.637	2.599	2.568	2.544	2.505	2.466	2.426	2.384	2.341	2.296	12		
2.554	2.515	2.484	2.459	2.420	2.380	2.339	2.297	2.252	2.206	13		
2.484	2.445	2.413	2.388	2.349	2.308	2.266	2.223	2.178	2.131	14		
2.424	2.385	2.352	2.326	2.294	2.261	2.224	2.182	2.140	2.096	15		
2.373	2.333	2.302	2.276	2.235	2.194	2.151	2.106	2.059	2.010	16		
2.329	2.289	2.257	2.230	2.190	2.148	2.104	2.058	2.011	1.960	17		
2.290	2.250	2.217	2.191	2.150	2.107	2.063	2.017	1.968	1.917	18		
2.256	2.215	2.182	2.155	2.114	2.071	2.026	1.980	1.930	1.878	19		
2.225	2.184	2.151	2.124	2.082	2.039	1.994	1.946	1.896	1.843	20		
2.197	2.156	2.123	2.096	2.054	2.010	1.965	1.916	1.866	1.812	21		
2.173	2.131	2.098	2.071	2.028	1.984	1.938	1.889	1.838	1.783	22		
2.150	2.109	2.075	2.048	2.005	1.961	1.914	1.865	1.813	1.757	23		
2.130	2.088	2.054	2.027	1.984	1.939	1.892	1.842	1.790	1.733	24		
2.111	2.069	2.035	2.007	1.964	1.919	1.872	1.822	1.768	1.711	25		
2.094	2.052	2.018	1.990	1.946	1.901	1.853	1.803	1.749	1.691	26		
2.078	2.036	2.002	1.974	1.930	1.884	1.836	1.785	1.731	1.672	27		
2.064	2.021	1.987	1.959	1.915	1.869	1.820	1.769	1.714	1.654	28		
2.050	2.007	1.973	1.945	1.901	1.854	1.806	1.754	1.698	1.638	29		
2.037	1.995	1.960	1.932	1.887	1.833	1.786	1.735	1.681	1.623	30		
2.026												

付表4 F分布表(パーセント点)

自由度が表にないときの $F_{\alpha}(\nu_1, \nu_2)$ は自由度の逆数による線形補間によって求める。

例3. 自由度 $\nu_1=80, \nu_2=5$ の上側 2.5 パーセント点は、 $\nu_1=80$ が表に含まれていないので、 $\nu_1=60$ に対する $F_{0.025}(60, 5)=6.123$ 、および $\nu_1=120$ に対する $F_{0.025}(120, 5)=6.069$ 、さらに、

$$\left(\frac{1}{80} - \frac{1}{120}\right) / \left(\frac{1}{60} - \frac{1}{120}\right) = 0.5$$

より、

$$F_{0.025}(4, 90)=6.069+(6.123-6.069)\times 0.5=6.096$$

とする。

$\alpha=0.025$

ν_1	1	2	3	4	5	6	7	8	9	10	12
ν_2	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28	968.63	976.71
1	38.506	39.000	39.165	39.248	39.298	39.331	39.355	39.373	39.387	39.398	39.415
2	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473	14.419	14.337
3	12.218	10.649	9.979	9.605	9.364	9.197	9.074	8.980	8.905	8.844	8.751
4	10.007	8.434	7.764	7.388	7.146	6.978	6.833	6.757	6.681	6.619	6.525
5	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461	5.366
6	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761	4.666
7	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295	4.200
8	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964	3.868
9	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.621
10	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526	3.430
11	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374	3.277
12	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250	3.153
13	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147	3.050
14	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.963
15	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986	2.889
16	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922	2.825
17	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866	2.769
18	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817	2.720
19	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.676
20	5.827	4.420	3.819	3.475	3.250	3.090	2.969	2.874	2.798	2.735	2.637
21	5.786	4.383	3.783	3.440	3.215	3.055	2.934	2.839	2.763	2.700	2.602
22	5.750	4.349	3.750	3.408	3.183	3.023	2.902	2.808	2.731	2.668	2.570
23	5.717	4.319	3.721	3.379	3.155	2.995	2.874	2.779	2.703	2.640	2.541
24	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677	2.613	2.515
25	5.659	4.265	3.670	3.329	3.105	2.945	2.824	2.729	2.653	2.590	2.491
26	5.633	4.242	3.647	3.307	3.083	2.923	2.802	2.707	2.631	2.568	2.469
27	5.610	4.221	3.626	3.286	3.063	2.903	2.782	2.687	2.611	2.547	2.448
28	5.588	4.201	3.607	3.267	3.044	2.884	2.763	2.669	2.592	2.529	2.430
29	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511	2.412
30	5.549	4.163	3.573	3.234	3.010	2.851	2.730	2.635	2.558	2.495	2.396
31	5.531	4.149	3.557	3.218	2.995	2.836	2.715	2.620	2.543	2.480	2.381
32	5.515	4.134	3.543	3.204	2.981	2.822	2.701	2.606	2.529	2.466	2.366
33	5.499	4.120	3.529	3.191	2.968	2.808	2.688	2.593	2.516	2.453	2.353
34	5.485	4.106	3.517	3.179	2.956	2.796	2.676	2.581	2.504	2.440	2.341
35	5.471	4.094	3.505	3.167	2.944	2.785	2.664	2.569	2.492	2.429	2.329
36	5.458	4.082	3.493	3.156	2.933	2.774	2.653	2.558	2.481	2.418	2.318
37	5.446	4.071	3.483	3.145	2.923	2.763	2.643	2.548	2.471	2.407	2.307
38	5.435	4.061	3.473	3.135	2.913	2.754	2.633	2.538	2.461	2.397	2.298
39	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388	2.288
40	5.377	4.009	3.422	3.086	2.864	2.705	2.584	2.489	2.412	2.348	2.248
50	5.340	3.975	3.390	3.054	2.833	2.674	2.553	2.458	2.381	2.317	2.216
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270	2.169
80	5.218	3.864	3.284	2.950	2.730	2.571	2.450	2.355	2.277	2.213	2.111
120	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222	2.157	2.055
co	5.024	3.689	3.116	2.786	2.567	2.408	2.288	2.192	2.114	2.048	1.945

14	16	18	20	24	30	40	60	120	∞	ν_1	ν_2
982.53	986.92	990.35	993.10	997.25	1001.4	1005.6	1009.8	1014.0	1018.3	1	
39.427	39.435	39.442	39.448	39.456	39.465	39.473	39.481	39.490	39.498	2	
14.277	14.232	14.196	14.167	14.124	14.081	14.037	13.992	13.947	13.902	3	
8.684	8.633	8.592	8.560	8.511	8.461	8.411	8.360	8.309	8.257	4	
6.456	6.403	6.362	6.329	6.278	6.227	6.175	6.123	6.069	6.015	5	
5.297	5.244	5.202	5.168	5.117	5.065	5.012	4.959	4.904	4.849	6	
4.596	4.543	4.501	4.467	4.415	4.362	4.309	4.254	4.199	4.142	7	
4.130	4.076	4.034	3.999	3.947	3.894	3.840	3.784	3.728	3.670	8	
3.798	3.744	3.701	3.667	3.614	3.560	3.505	3.449	3.392	3.333	9	
3.550	3.496	3.453	3.419	3.365	3.311	3.255	3.198	3.140	3.080	10	
3.359	3.304	3.261	3.226	3.173	3.118	3.061	3.004	2.944	2.883	11	
3.206	3.152	3.108	3.073	3.019	2.963	2.906	2.848	2.787	2.725	12	
3.082	3.027	2.983	2.948	2.893	2.837	2.780	2.720	2.659	2.595	13	
2.979	2.923	2.879	2.844	2.789	2.732	2.674	2.614	2.552	2.487	14	
2.891	2.836	2.792	2.756	2.701	2.644	2.585	2.524	2.461	2.395	15	
2.817	2.761	2.717	2.681	2.623	2.568	2.509	2.447	2.383	2.316	16	
2.753	2.697	2.652	2.616	2.560	2.502	2.442	2.380	2.315	2.247	17	
2.696	2.640	2.596	2.559	2.503	2.445	2.384	2.321	2.256	2.187	18	
2.647	2.591	2.546	2.509	2.452	2.394	2.333	2.278	2.208	2.143	19	
2.603	2.547	2.501	2.464	2.408	2.349	2.287	2.223	2.156	2.085	20	
2.564	2.507	2.462	2.425	2.368	2.308	2.246	2.182	2.114	2.042	21	
2.528	2.472	2.426	2.389	2.331	2.272	2.210	2.145	2.076	2.003	22	
2.497	2.440	2.394	2.357	2.299	2.239	2.176	2.111	2.041	1.968	23	
2.468	2.411	2.365	2.327	2.269	2.209	2.146	2.080	2.010	1.935	24	
2.441	2.384	2.338	2.300	2.242	2.182	2.118	2.052	1.981	1.906	25	
2.417	2.360	2.314	2.276	2.217	2.157	2.093	2.026	1.954	1.878	26	
2.395	2.337	2.291	2.253	2.195	2.133	2.069	2.002	1.930	1.853	27	
2.374	2.317	2.270	2.232	2.174	2.112	2.048	1.980	1.907	1.829	28	
2.355	2.298	2.251	2.213	2.154	2.092	2.028	1.959	1.886	1.807	29	
2.338	2.280	2.233	2.195	2.136	2.074	2.009	1.940	1.866	1.787	30	
2.321	2.263	2.217	2.178	2.119	2.057	1.991	1.922	1.848	1.768	31	
2.306	2.248	2.201	2.163	2.103	2.041	1.975	1.905	1.831	1.750	32	
2.292	2.234	2.187	2.148	2.088	2.026	1.960	1.890	1.815	1.733	33	
2.278	2.220	2.173	2.135	2.075	2.012	1.946	1.875	1.799	1.717	34	
2.266	2.207	2.160	2.122	2.062	1.999	1.932	1.861	1.785	1.702	35	
2.254	2.196	2.148	2.110	2.049	1.986	1.919	1.848	1.772	1.687	36	
2.243	2.184	2.137	2.098</								

付表4 F分布表(パーセント点)

自由度が表にないときの $F_{\alpha}(\nu_1, \nu_2)$ は自由度の逆数による線形補間によって求める。

例4. 自由度 $\nu_1=80$, $\nu_2=90$ の上側1パーセント点は, $\nu_1=80$ および $\nu_2=90$ がともに表に含まれていないので, 2段階に補間をおこなう。最初に $(\nu_1, \nu_2)=(80, 80)$ および $(80, 120)$ に対する値を, $\nu_1=60$ と $\nu_1=120$ の値を用いて求める。すなわち

$$\left(\frac{1}{80} - \frac{1}{120}\right) / \left(\frac{1}{60} - \frac{1}{120}\right) = 0.5$$

$\alpha=0.01$

ν_1	1	2	3	4	5	6	7	8	9	10	12
ν_2	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5	6055.8	6106.3
1	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399	99.416
2	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229	27.052
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546	14.374
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051	9.888
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874	7.718
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620	6.469
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814	5.667
9	10.561	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257	5.111
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849	4.706
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539	4.397
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296	4.155
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100	3.960
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939	3.800
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805	3.666
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691	3.553
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593	3.455
18	8.285	6.013	5.092	4.579	4.248	4.013	3.841	3.705	3.597	3.508	3.371
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.531	3.523	3.434	3.297
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368	3.231
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398	3.310	3.173
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258	3.121
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299	3.211	3.074
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168	3.032
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217	3.129	2.993
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094	2.958
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149	3.062	2.926
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032	2.896
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092	3.005	2.868
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979	2.843
31	7.530	5.362	4.484	3.993	3.675	3.449	3.281	3.149	3.043	2.955	2.820
32	7.499	5.336	4.459	3.969	3.652	3.427	3.258	3.127	3.021	2.934	2.798
33	7.471	5.312	4.437	3.948	3.630	3.406	3.238	3.106	3.000	2.913	2.777
34	7.444	5.289	4.416	3.927	3.611	3.386	3.218	3.087	2.981	2.894	2.758
35	7.419	5.268	4.396	3.908	3.592	3.368	3.200	3.069	2.963	2.876	2.740
36	7.396	5.248	4.377	3.890	3.574	3.351	3.183	3.052	2.946	2.859	2.723
37	7.373	5.229	4.360	3.873	3.558	3.334	3.167	3.036	2.930	2.843	2.707
38	7.353	5.211	4.343	3.858	3.542	3.319	3.152	3.021	2.915	2.828	2.692
39	7.333	5.194	4.327	3.843	3.528	3.305	3.137	3.006	2.901	2.814	2.678
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801	2.665
45	7.234	5.110	4.249	3.767	3.454	3.232	3.066	2.935	2.830	2.743	2.608
50	7.171	5.057	4.199	3.720	3.408	3.186	3.020	2.890	2.785	2.698	2.562
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718	2.632	2.496
80	6.963	4.881	4.036	3.563	3.255	3.036	2.871	2.742	2.637	2.551	2.415
120	6.851	4.787	3.949	3.480	3.174	2.956	2.792	2.663	2.559	2.472	2.336
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.639	2.511	2.407	2.321	2.185

より,

$$F_{0.01}(80, 80) = 1.630 + (1.746 - 1.630) \times 0.5 = 1.688$$

$$F_{0.01}(80, 120) = 1.533 + (1.656 - 1.533) \times 0.5 = 1.595$$

を得る。次に,

$$\left(\frac{1}{90} - \frac{1}{120}\right) / \left(\frac{1}{80} - \frac{1}{120}\right) = 0.067$$

より,

$$F_{0.01}(80, 90) = 1.595 + (1.688 - 1.595) \times 0.667 = 1.657$$

とする。

ν_1	14	16	18	20	24	30	40	60	120	∞	ν_2
6142.7	6170.1	6191.5	6208.7	6234.6	6260.6	6286.8	6313.0	6339.4	6365.9	1	1
99.428	99.437	99.444	99.449	99.458	99.466	99.474	99.482	99.491	99.499	2	2
26.924	26.827	26.751	26.690	26.598	26.505	26.411	26.316	26.221	26.125	3	3
14.249	14.154	14.080	14.020	13.929	13.838	13.745	13.652	13.558	13.463	4	4
9.770	9.680	9.610	9.553	9.466	9.379	9.291	9.202	9.112	9.020	5	5
7.605	7.519	7.451	7.396	7.313	7.229	7.143	7.057	6.969	6.880	6	6
6.359	6.275	6.209	6.155	6.074	5.992	5.908	5.824	5.737	5.650	7	7
5.559	5.477	5.412	5.359	5.279	5.198	5.116	5.032	4.946	4.859	8	8
5.005	4.924	4.860	4.808	4.729	4.649	4.567	4.483	4.398	4.311	9	9
4.601	4.520	4.457	4.405	4.327	4.247	4.165	4.082	3.996	3.909	10	10
4.293	4.213	4.150	4.099	4.021	3.941	3.860	3.776	3.690	3.602	11	11
4.052	3.972	3.909	3.858	3.780	3.701	3.619	3.535	3.449	3.361	12	12
3.857	3.778	3.716	3.665	3.587	3.507	3.425	3.341	3.255	3.165	13	13
3.698	3.619	3.556	3.505	3.427	3.348	3.266	3.181	3.094	3.004	14	14
3.564	3.485	3.423	3.372	3.294	3.214	3.132	3.053	2.971	2.888	15	15
3.451	3.372	3.310	3.259	3.181	3.101	3.018	2.933	2.845	2.753	16	16
3.353	3.275	3.212	3.162	3.084	3.003	2.920	2.835	2.746	2.653	17	17
3.269	3.190	3.128	3.077	2.999	2.919	2.835	2.749	2.660	2.566	18	18
3.195	3.116	3.054	3.003	2.925	2.844	2.761	2.674	2.584	2.489	19	19
3.130	3.051	2.989	2.938	2.859	2.778	2.695	2.608	2.517	2.421	20	20
3.072	2.993	2.931	2.880	2.801	2.720	2.636	2.548	2.457	2.360	21	21
3.019	2.941	2.879	2.827	2.749	2.667	2.583	2.495	2.403	2.305	22	22
2.973	2.894	2.832	2.781	2.702	2.620	2.535	2.447	2.354	2.256	23	23
2.930	2.852	2.789	2.738	2.659	2.577	2.492	2.403	2.310	2.211	24	24
2.892	2.813	2.751	2.699	2.620	2.538	2.453	2.364	2.270	2.169	25	25
2.857	2.778	2.715	2.664	2.585	2.503	2.417	2.327	2.233	2.131	26	26
2.824	2.746	2.683	2.632	2.552	2.470	2.384	2.294	2.198	2.097	27	27
2.795	2.716	2.653	2.602	2.522	2.440	2.354	2.263	2.167	2.064	28	28
2.767	2.689	2.626	2.574	2.495	2.412	2.325	2.234	2.138	2.034	29	29
2.747	2.677	2.616	2.549	2.469	2.386	2.299	2.208	2.111	2.006	30	30
2.718	2.640	2.577	2.525	2.445	2.362	2.27					

付表4 F分布表(パーセント点)

自由度が表にないときは近似式を用いることもできる。少し複雑だが、ポールソン-竹内の近似式

$$F_\alpha(\nu_1, \nu_2) = \left[\frac{(1-a)(1-b) + z_{\alpha} \sqrt{(1-a)^2 b + (1-b)^2 a - ab z_{\alpha}^2}}{(1-b)^2 - bz_{\alpha}^2} \right]^{\nu_2} \quad \text{ただし, } a=2/(9\nu_1), \quad b=2/(9\nu_2)$$

はよい精度をもつ。ここで z_{α} は標準正規分布の上側 100α パーセント点である。

例5. $\nu_1=30$, $\nu_2=40$, $\alpha=0.005$ に対しては、表から $F_{0.005}(30, 40)=2.401$ が得られる。ポールソン-竹内

$\alpha=0.005$

ν_1	1	2	3	4	5	6	7	8	9	10	12
ν_2	16210.7	19999.3	21614.7	22499.6	23055.8	23437.1	23714.6	23923.4	24091.0	24224.3	24426.4
2	198.50	199.00	199.17	199.25	199.30	199.33	199.36	199.37	199.39	199.40	199.42
3	55.552	49.799	47.467	46.195	43.392	44.838	44.434	44.126	43.882	43.686	43.387
4	31.333	26.284	24.259	23.155	22.456	21.975	21.622	21.352	21.139	20.967	20.705
5	22.783	18.314	16.530	15.356	14.940	14.513	14.200	13.961	13.772	13.618	13.384
6	18.635	14.544	12.917	12.028	11.464	11.073	10.786	10.566	10.391	10.230	10.034
7	16.236	12.404	10.882	10.050	9.522	9.155	8.885	8.678	8.514	8.380	8.176
8	14.688	11.042	9.596	8.805	8.302	7.932	7.694	7.496	7.339	7.211	7.015
9	13.614	10.107	8.717	7.956	7.471	7.134	6.885	6.693	6.541	6.417	6.227
10	12.826	9.427	8.081	7.343	6.872	6.345	6.302	6.116	5.968	5.847	5.661
11	12.226	8.912	7.600	6.881	6.422	6.102	5.865	5.682	5.537	5.418	5.236
12	11.754	8.510	7.226	6.521	6.071	5.757	5.525	5.345	5.202	5.085	4.906
13	11.374	8.186	6.926	6.233	5.791	5.482	5.253	5.076	4.933	4.820	4.643
14	11.060	7.922	6.680	5.998	5.562	5.257	5.031	4.857	4.717	4.603	4.428
15	10.798	7.701	6.476	5.803	5.372	5.071	4.847	4.674	4.536	4.424	4.250
16	10.375	7.514	6.303	5.638	5.212	4.913	4.692	4.521	4.384	4.272	4.099
17	10.384	7.354	6.156	5.497	5.075	4.779	4.539	4.389	4.254	4.142	3.971
18	10.218	7.215	6.028	5.375	4.956	4.663	4.445	4.276	4.141	4.030	3.860
19	10.073	7.093	5.916	5.268	4.853	4.561	4.345	4.177	4.043	3.933	3.763
20	9.944	6.986	5.818	5.174	4.762	4.472	4.257	4.090	3.956	3.847	3.678
21	9.830	6.891	5.730	5.091	4.681	4.393	4.179	4.013	3.880	3.771	3.602
22	9.727	6.806	5.652	5.017	4.609	4.322	4.109	3.944	3.812	3.703	3.535
23	9.635	6.730	5.582	4.950	4.544	4.259	4.047	3.882	3.750	3.642	3.473
24	9.551	6.661	5.519	4.890	4.486	4.202	3.991	3.826	3.693	3.587	3.420
25	9.475	6.598	5.462	4.833	4.433	4.150	3.939	3.776	3.645	3.537	3.370
26	9.406	6.541	5.409	4.783	4.384	4.103	3.893	3.730	3.599	3.492	3.323
27	9.342	6.489	5.361	4.740	4.340	4.059	3.850	3.687	3.557	3.450	3.284
28	9.284	6.440	5.317	4.698	4.300	4.020	3.811	3.649	3.519	3.412	3.246
29	9.230	6.396	5.276	4.659	4.262	3.983	3.775	3.613	3.483	3.377	3.211
30	9.180	6.355	5.239	4.623	4.228	3.949	3.742	3.580	3.450	3.344	3.179
31	9.133	6.317	5.204	4.590	4.196	3.918	3.711	3.549	3.420	3.314	3.149
32	9.090	6.281	5.171	4.559	4.166	3.889	3.682	3.521	3.392	3.286	3.121
33	9.050	6.248	5.141	4.531	4.138	3.861	3.655	3.495	3.366	3.260	3.095
34	9.012	6.217	5.113	4.504	4.112	3.836	3.630	3.470	3.341	3.235	3.071
35	8.976	6.188	5.086	4.479	4.088	3.812	3.607	3.447	3.318	3.212	3.048
36	8.943	6.161	5.062	4.455	4.065	3.790	3.585	3.425	3.296	3.191	3.027
37	8.912	6.135	5.038	4.433	4.043	3.769	3.564	3.404	3.276	3.171	3.007
38	8.882	6.111	5.016	4.412	4.023	3.749	3.545	3.385	3.257	3.152	2.988
39	8.854	6.088	4.995	4.392	4.004	3.731	3.526	3.367	3.239	3.134	2.970
40	8.828	6.066	4.976	4.374	4.088	3.713	3.509	3.350	3.222	3.117	2.953
45	8.715	5.974	4.892	4.294	3.909	3.638	3.435	3.276	3.149	3.044	2.881
50	8.626	5.902	4.826	4.232	3.849	3.579	3.376	3.219	3.092	2.988	2.825
60	8.495	5.795	4.729	4.140	3.760	3.492	3.291	3.134	3.008	2.904	2.742
80	8.335	5.665	4.611	4.029	3.652	3.387	3.188	3.032	2.907	2.803	2.641
120	8.179	5.339	4.497	3.921	3.548	3.285	3.087	2.933	2.808	2.705	2.544
∞	7.879	5.298	4.279	3.715	3.350	3.091	2.897	2.744	2.621	2.519	2.358

ン-竹内の近似式を用いると、

$$a=2/(9 \times 30)=0.0074074, \quad b=2/(9 \times 40)=0.0055556, \quad z_{0.005}=2.5758$$

から

$$F_{0.005}(30, 40) \approx 2.404$$

となる。

ν_1	14	16	18	20	24	30	40	60	120	∞	ν_2
24571.8	24681.3	24767.2	24836.0	24939.6	25043.6	25148.2	25253.1	25358.6	25464.5	25464.5	1
199.43	199.44	199.44	199.45	199.46	199.47	199.47	199.48	199.49	199.50	199.50	2
43.172	43.008	42.880	42.778	42.622	42.466	42.308	42.149	41.989	41.828	41.828	3
20.515	20.371	20.258	20.167	20.030	19.892	19.752	19.611	19.468	19.325	19.325	4
13.215	13.086	12.985	12.903	12.780	12.656	12.530	12.402	12.274	12.144	12.144	5
9.877	9.758	9.664	9.589	9.474	9.358	9.241	9.122	9.001	8.879	8.879	6
8.028	7.915	7.826	7.734	7.645	7.534	7.422	7.309	7.193	7.076	7.076	7
6.872	6.763	6.678	6.608	6.503	6.396	6.288	6.177	6.065	5.951	5.951	8
6.089	5.983	5.899	5.832	5.729	5.625	5.519	5.410	5.300	5.188	5.188	9
5.326	5.422	5.340	5.274	5.173	5.071	4.966	4.859	4.750	4.639	4.639	10
5.103	5.001	4.921	4.855	4.756	4.654	4.551	4.445	4.337	4.226	4.226	11
4.775	4.674	4.595	4.530	4.431	4.331	4.228	4.123	4.015	3.904	3.904	12
4.513	4.413	4.334	4.270	4.173	4.073	3.970	3.866	3.758	3.647	3.647	13
4.299	4.200	4.122	4.059	3.961	3.862	3.760	3.655	3.547	3.436	3.436	14
4.122	4.024	3.946	3.883	3.786	3.687	3.585	3.480	3.372	3.260	3.260	15
3.972	3.875	3.795	3.714	3.638	3.559	3.437	3.332	3.224	3.112	3.112	16
3.844	3.747	3.670	3.607	3.511	3.412	3.311	3.206	3.097	2.984	2.984	17
3.734	3.637	3.560	3.498	3.402	3.303	3.201	3.096	2.987	2.873	2.873	18
3.638	3.541	3.465	3.402	3.306	3.208	3.106	3.000	2.891	2.776	2.776	19
3.553	3.457	3.380	3.318	3.222	3.123	3.022	2.916	2.806	2.690	2.690	20
3.478	3.382	3.305	3.243	3.147	3.049	2.947	2.841	2.730	2.614	2.614	21
3.411	3.315	3.239	3.176	3.081	2.982	2.880	2.774	2.663	2.545	2.545	22
3.351	3.255	3.179	3.116	3.021	2.922	2.820	2.713	2.602	2.484	2.484	23
3.296	3.201	3.125	3.062	2.967	2.868	2.765	2.658	2.546	2.428	2.428	24
3.247	3.151	3.075	3.013	2.918	2.819	2.716	2.609	2.496	2.377	2.377	25
3.202	3.107	3.031	2.968	2.873	2.774	2.671	2.563	2.450	2.330	2.330	26
3.161	3.066	2.990	2.928	2.832	2.733	2.630	2.522	2.408	2.287	2.287	27
3.123	3.028	2.952	2.890	2.794	2.695	2.592	2.483	2.369	2.247	2.247	28
3.088	2.993	2.917	2.855	2.759	2.660	2.557	2.448	2.333	2.210	2.210	29
3.036	2.961	2.885	2.823	2.727	2.628	2.524	2.415	2.300	2.176	2.176	30
3.026	2.931	2.855	2.793	2.707	2.598	2.494	2.385	2.269	2.144	2.144	31
2.998	2.904	2.828	2.766	2.670	2.570</						

練習問題の解答

パソコンによるデータ分析の実習例としてふさわしいものについて、一部くわしい解答を与えていないものがある。

第1章

1.1 $E(X) = \sum_{r=0}^n r_n C_r p^r q^{n-r} = np \sum_{r=1}^n n-1 C_{r-1} p^{r-1} q^{n-r} = np \sum_{r=0}^{n-1} n-1 C_r p^r q^{(n-1)-r} = np(p+q)^{n-1} = np$ 。同様に $E(X(X-1)) = n(n-1)p^2$ が示せる。したがって、 $V(X) = E(X(X-1)) + E(X) - \{E(X)\}^2 = npq$

1.2 直接計算でももとまるが、ここではモーメント母関数を利用する。 $M(t) = \sum_{r=1}^{\infty} e^{tr} p q^{r-1} = p e^t / (1 - q e^t)$ 。したがって、 $M^{(1)}(0) = 1/p$

1.3 $Y_i (i=1, 2, \dots, k)$ を互いに独立で、パラメータ p の同一の幾何分布に従うとすれば、 $X = Y_1 + Y_2 + \dots + Y_k$ と表現できる。したがって、問題 1.2 より $E(X) = k/p$

1.4 i) 等式 $(1+x)^m (1+x)^n = (1+x)^{m+n}$ の両辺を展開して x^1 の係数を比較すれば、 $\sum_{i=0}^m C_{i,n} C_{1-i} = {}_{c+n} C_1$ が導ける。ただし、 ${}_a C_b = 0$ ($b > a$ あるいは $b < 0$) とする。ii) まず $p_r = {}_n C_{r-n} C_r / {}_n C_R$ に注意する。 N と R が、 n と r に比較して十分大きいとき、本式より $p_r = {}_n C_r R(R-1)\cdots(R-r+1)(N-R)\cdots(N-R-n+r+1)/(N(N-1)\cdots(N-n+1)) = {}_n C_r p^r q^{n-r}$ となる。

1.5 $E(X) = e^{-\lambda} \sum_{k=0}^{\infty} k \lambda^k / k! = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \lambda^{k-1} / (k-1)! = \lambda e^{-\lambda} e^{\lambda} = \lambda$ 。同様に $E(X(X-1)) = \lambda^2$ 。したがって、 $V(X) = \lambda$

1.6 i) $\lambda = 0.61$ を用いて $200 \times e^{-\lambda} \lambda^k / k!$ を、各 k に対して計算する。ii) i) に比べてはまりは良くない。ポアソン分布は、ベルヌーイ試行の極限である。交通事故の確率は天候などに左右される。また、玉突事故などでは多数の死者が出る。したがって、ベルヌーイ試行が満たすべき 3 条件、独立性、定常性、二値性の成立が疑わしい。

1.7 i) $(a+b)/2$ ii) $(b-a)^2/12$

1.8 (1.9) と (1.12) より明らか。

1.9 i) S_1 のモーメント母関数は $\lambda / (\lambda - t)$ ($t < \lambda$)。 $t > \lambda$ のときは存在しない。したがって、(1.12) より $M(t) = \{\lambda / (\lambda - t)\}^n$ ($t < \lambda$)。 $t > \lambda$ のときは存在しない。ii) $M(t) = \{\lambda / (\lambda - t)\}^n$ ($t < \lambda$)。 $t > \lambda$ のときは存在しない。iii) ii) で $s=n$ とおけばよい。

1.10 $S_n = X_1 + X_2 + \dots + X_n$ とおけば、問題 1.8 と数学的帰納法により、 S_n はパラメータ n のポアソン分布に従う。一方、 $e^{-n} \sum_{k=0}^n n^k / k! = P(S_n \leq n) = P((S_n - n)/n \leq 0)$ であるから、右辺に中心極限定理を適用すればよい。

第2章

2.1 略。

2.2 行列 X は以下の通り。 $X'X$ は略。 $X'y$ は自ら計算せよ。

$$\begin{array}{lll} \text{i)} \quad X = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & & 1 \\ \vdots & & \\ 1 & & 1 \end{bmatrix}^{15} & \text{ii)} \quad X = \begin{bmatrix} 1 & 1 & & O \\ 1 & & 1 & 5 \\ \vdots & & 1 & \\ & & 1 & 5 \\ \vdots & & 1 & \\ & & 1 & 5 \\ 1 & O & & \end{bmatrix} & \text{iii)} \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \\ 1 & 7 & 49 \\ 1 & 8 & 64 \\ 1 & 9 & 81 \\ 1 & 10 & 100 \end{bmatrix} \end{array}$$

2.3 自ら試みよ。

2.4 i) すべて等しく, a) $n=11$ b) $\bar{x}=90$ c) $\bar{y}=7.5$ d) $r_{xy}=0.816$ e) $\hat{\beta}_1=0.5$ f) $y=3+0.5x$ g) $S_{xx}=111.0$ h) $S_{xy}=55.25$ i) $S_{yy}^2/S_{xx}=27.50$ j) $t=4$.

24 ii) 54 頁記事参照。

2.5 i) $s.e.=\sqrt{930.825}/30.51$ ii)～v) 自ら試みよ。

2.6 $S_{xx}=924$, $S_{yy}=1,222$, $S_{xy}=-936$. よって, $\hat{\beta}_1=-936/924=-1.013$, $\hat{\beta}_0=64$.

247. 回帰直線 $y=64.247-1.013x$. また $\delta^2=(1222-(-936)^2/924)/10=27.38$. $t=-1.013/\sqrt{27.38/924}=-5.88$ (自由度=10). 傾き=0 の仮説は棄却される。

2.7 i) $\bar{x}=6.167$, $\bar{y}=9.000$, $S_{xx}=21.853$, $S_{xy}=41.520$. したがって, $y=1.900x-2.716$. なお, $S_{yy}=97.400$ より $S_e=18.515$, $\delta=1.193$. ii) 10.39. iii) 9.61, 11.17, iv) $t=0.392$ で有意でない。

2.8 i) $y_{(A)}=20.014+0.942x_{(A)}$, $S_{e(A)}=1.663$, $\delta_{(A)}^2=0.4157$, $y_{(B)}=21.940+0.868x_{(B)}$, $S_{e(B)}=5.710$, $\delta_{(B)}^2=1.4276$. ii) $\delta_{(B)}^2/\delta_{(A)}^2=3.43 < F_{0.025}(4,4)=9.605$ だから仮説は棄却されない。iii) 共通の分散の推定値は $\delta^2=(1.663+5.710)/8=0.9216$. また, $S_{xx(A)}=39.333$, $S_{xx(B)}=43.333$. 仮説のもとで, $t=(\hat{\beta}_{1(A)}-\hat{\beta}_{1(B)})/\sqrt{\delta^2(1/S_{xx(A)}+1/S_{xx(B)})}$ は $t(8)$ に従う。いま, $t=(0.942-0.868)/\sqrt{0.9216(1/39.333+1/43.333)}=0.074/0.2114=0.35$ で, $|t| < t_{0.025}(8)=2.306$ だから仮説は棄却されない。iv) 仮説のもとで, $t=(\hat{\beta}_{0(A)}-\hat{\beta}_{0(B)})/\sqrt{\delta^2(2/n+\bar{x}_{(A)}^2/S_{xx(A)}+\bar{x}_{(B)}^2/S_{xx(B)})}$ は $t(8)$ に従う。いま, $t=(20.014-21.940)/\sqrt{0.9216(2/8+8.33^2/39.333+8.667^2/43.333)}=1.926/1.943=0.99$ で, $|t| < t_{0.025}(8)=2.306$ だから仮説は棄却されない。v) 2直線が異なるとしたときの残差平方和は $S_e=S_{e(A)}+S_{e(B)}=1.663+5.710=7.373$, また, 2つのデータに単一の直線をあてはめると, $y=20.853+0.919x$ となり, 残差平方和は $S_e'=12.526$. 仮説のもとで $F=((S_e'-S_e)/2)/(S_e/8)$ は $F(2,8)$ に従う。いま, $F=2.80 < F_{0.05}(2,8)=4.459$ だから仮説は棄却されない。

2.9 i) 正規方程式は, $\begin{bmatrix} 6 & 0 & 50 \\ 0 & 6 & 52 \\ 50 & 52 & 950 \end{bmatrix} \begin{bmatrix} \beta_{0(A)} \\ \beta_{0(B)} \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 167.2 \\ 176.8 \\ 3000.3 \end{bmatrix}$ で, 2直線はそれぞれ $y_{(A)}=20.336+0.904x_{(A)}$, $y_{(B)}=21.635+0.904x_{(B)}$. ii) $S_e=7.486$, したがって, $\delta^2=7.486/9=0$.

832. iii) 2つのデータに単一の直線をあてはめたときの残差平方和は $S_e'=12.526$. 仮説のもとで $F=(S_e'-S_e)/(S_e/9)$ は $F(1,9)$ に従う。いま, $F=6.06 > F_{0.05}(1,9)=5.117$ だから仮説は棄却される。

2.10 i) $\hat{\beta}=\sum x_i y_i / \sum x_i^2$. ii) 正規分布 $N(\beta, \sigma^2 / \sum x_i^2)$. iii) $\sigma^2 = S_e / (n-1)$. ただし, $S_e = \sum y_i^2 - (\sum x_i y_i)^2 / \sum x_i^2$. iv) $\hat{\beta} \pm t_{0.025}(n-1) \delta \sqrt{x^2 / \sum x_i^2}$.

第3章

3.1 i) $\bar{y}_1=0.8524$, $\delta_1^2=0.0071$, $\bar{y}_2=1.0918$, $\delta_2^2=0.0170$ より, $F=0.4188$. $1/F_{0.05}(8,8)=1/4.433=0.2256$ だから, 有意水準 5% では有意ではない。ii) $|t|=|-4.62| < t_{0.05}(16)=2.120$ だから, 平均には有意差がある。iii) $[-0.349, -0.129]$.

3.2 分散分析表は次のとおり。グラフは略。

要因	平方和	自由度	平均平方	F 値
樹脂濃度(A)	62.75	3	20.92	6.44**
誤差	39.00	12	3.25	
計	101.75	15		

3.3 略。

3.4 略。

3.5 $\bar{y}_{ij..} - y_{ij..} \pm t_{0.025}(\nu_e) \sqrt{2V_e/r}$. $(\bar{y}_{i..} + \bar{y}_{j..}) - (\bar{y}_{i..} + \bar{y}_{j..}) \pm t_{0.025}(\nu_e) \sqrt{2V_e(a+b)/abr}$.

3.6 分散分析表は次のとおり。

要因	平方和	自由度	平均平方	F 値
反応温度(A)	763.0	2	381.5	19.78**
触媒量(B)	339.5	3	113.2	5.87*
交互作用(A×B)	25.7	6	4.28	0.22
誤差	231.5	12	19.29	
計	1359.6	23		

第4章

4.1 自ら試みよ。

4.2 i) $\log L(p)=n \log p + (T-n) \log(1-p)$ を微分せよ。 $\bar{p}=100/192=0.521$. ii) $I_1(p)=1/p^2(1-p)$. 検定の棄却域は $z=\sqrt{n}(\bar{p}-p_0)/p_0\sqrt{1-p_0} > z_\alpha$. このデータでは $z=0.594 < 1.645$ で仮説は棄却されない。iii) $z=(p_0 T - n)/\sqrt{n(1-p_0)} < -z_\alpha$. このデータでは $z=-0.566 < -1.645$ で仮説は棄却されない。iv) $w=2 \log(L(\bar{p})/L(p_0))=2n \log(n/Tp_0)+2(T-n) \log((T-n)/T(1-p_0)) > \chi_{\alpha}^2(1)$. このデータでは $w=0.333 < 3.84$ で仮説は棄却されない。v) ii) と iii) の z の値はよく似ている。ii) と iii) が両側検定なら $z^2 > \chi_{\alpha}^2(1)$ が棄却域となるが、これらの z^2 と iii) の w の値もよく似ている。

- 4.3 i) $\log L(\sigma) = \frac{n}{2} \log(2\pi) - n \log \sigma - \sum x_i^2 / 2\sigma^2$ から $\bar{\sigma} = \sqrt{\sum x_i^2 / n}$. ii) $\sigma^2 / 2n$.
- 4.4 故障時間の合計を S とすると, $\log L(\lambda) = n \log \lambda - \lambda S$ より $\bar{\lambda} = n/S$. このデータでは $\bar{\lambda} = 10/6629 = 0.00151 [\text{min}^{-1}]$ ii) $\exp(-t_0 \bar{\lambda}) = 0.22$

4.5 $\bar{\theta} = \max x_i$.

第5章

5.1 略.

5.2 $\chi^2 = 13.51 < \chi^2_{0.05}(13) = 22.36$ だから仮説は棄却されない.

5.3 $\chi^2 = 50.47 > \chi^2_{0.05}(3) = 7.815$ だから仮説は棄却される.

5.4 (i) 略. (ii) $\chi^2 = 0.933 < \chi^2_{0.05}(3) = 7.815$ だから仮説は棄却されない.

5.5 (i) $\bar{x} = 0.477$, $s^2 = 0.9355$. (ii) $\chi^2 = 131.62 > \chi^2_{0.05}(3) = 7.815$ だからポアソン分布はあてはまらない. (iii) $\hat{p} = 0.478$, $\hat{a} = 0.409$, $\chi^2 = 5.94 < \chi^2_{0.05}(2) = 5.99$ だから負の二項分布の仮説は棄却されない.

5.6 $\theta_1 = 20.631$, $\theta_2 = 19.079$, $\theta_3 = 17.197$, $\theta_4 = 16.756$, $\theta_5 = 15.910$, $\theta_6 = 10.428$. $\chi^2 = 6.84 < \chi^2_{0.05}(10) = 18.307$ だからモデルは成立するものとみなされる.

5.7 10⁴ 行までに対する(一様分布への)適合度の χ^2 合成度統計量を χ_i^2 とおくと, $\chi_1^2 = 9.32$, $\chi_5^2 = 4.09$, $\chi_6^2 = 5.51$, $\chi_7^2 = 2.78$, $\chi_8^2 = 7.27$, $\chi_9^2 = 4.92$. 一方, $\chi^2_{0.05}(9) = 16.92$. いずれも, 円周率のこれらの桁で, 0~9 の数字は等確率(一様分布)で分布するという帰無仮説は棄却されない.

第6章

6.1 表の回数を X として, $X=0, 4$ のとき棄却する検定方式 δ の検出力関数は $\beta_\delta(p) = P_p(X=4) + P_p(X=0) = p^4 + (1-p)^4$. グラフは略.

6.2 $\tau = \sqrt{5 \cdot 5 / (5+5)} \cdot 1/1 = \sqrt{10} = 1.581$. 次に $\alpha = 0.05$ なら $z_{0.05} = 1.645$, $z_{0.025} = 1.96$. ゆえに片側検定, 両側検定に対応して, それぞれ $1 - \Phi(1.645 - 1.581) = 0.474$, $1 - \Phi(1.96 - 1.581) - \Phi(-1.96 - 1.581) = 0.352$.

6.3 右片側検定なら, $\beta_\delta(\mu_1, \mu_2) = P(Z > z_\alpha) = P\left(\sqrt{\frac{mn}{m+n}} \cdot \frac{(\bar{X} - \mu_1) + (\bar{Y} - \mu_2)}{\sigma_0} > z_\alpha - \tau\right) = 1 - \Phi(z_\alpha - \tau)$. 第3式の()内左辺は標準化されて, 標準正規分布に従っていることに注意. 両側も同様.

6.4 (6.29) で $\beta_\delta(p) = 0.9$ を解く. (6.32) を求めた場合に習う. $p_0 = 0.4$, $p = 0.6$, $\alpha = 0.05$ とし, $z_{0.9} = -1.28$, $\sqrt{p_0(1-p_0)} = \sqrt{p(1-p)} = 0.49$, $p - p_0 = 0.2$ だから $-1.28 = \{(1.645)(0.49) - \sqrt{n} \cdot 0.2\} / 0.49$, よって $n = 51.4$. 少くとも 52.

6.5 $H_0: \mu = \mu_0$ 対 $H: \mu = \mu_1 (\mu_1 > \mu_0)$ の最強力検定は, ネイマン・ピアソンの補題により

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(x_i - \mu_1)^2/2\sigma_0^2} / \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(x_i - \mu_0)^2/2\sigma_0^2} > (\leq) k$$

で, H_0 の棄却(採択)を決める. この不等式を整理すると検定方式は $\bar{x} > C$. (なお, $C = (\mu_1 - \mu_0)/2 + (\sigma_0^2 \log k) / [n(\mu_1 - \mu_0)]$ となる.). したがって, C を有意水準が α となるように, $\sqrt{n}(C - \mu_0)/\sigma_0 = z_\alpha$ とすれば, 検定は $Z > z_\alpha$ となるが, この C の決め方は対立仮説 μ_1 に依らず, 検定はあらゆる対立仮説 μ_1 (ただし, $> \mu_0$)に対し最強力となる.

7.1 i) $I(\theta) = 1/2$, $V(\theta_{\text{ML}}) = 2/n$. ii) $V(X_{\text{med}}) = \pi^2/4 n = 2.467/n$. iii) $V(X_{\text{med}})/V(\hat{\theta}_{\text{ML}}) = 1.234$.

7.2 省略.

7.3 $\hat{\theta}_{\text{ML}} = 8.43$.

7.4 i) 省略. ii) $\bar{X} = 125.69/25 = 5.028$. iii) $X_{\text{med}} = 5.02$. iv) $\hat{\theta}_{\text{trim}}(3) = 95.24/19 = 5.013$. v) $\hat{\theta}_{\text{H}} = (70.45 + 0.1525(5-6))/14 = 5.021$. vi) 省略.

7.5 $n = 17$, $N = 12$ だから, 仮説は棄却されない.

7.6 第2標本の順位和 $W = 158$. 表より下側 2.5 パーセント点は 94 だから, 上側 2.5 パーセント点は $10(10+15+1) - 94 = 166$. したがって, 仮説は棄却されない.
 7.7 i) 省略. ii) (A) $b_1 = 0.80$, $b_2 = -0.40$ だから正規分布の仮説は棄却されない.
 (B) $b_1 = 0.90$, $b_2 = 2.40$ だから正規分布の仮説は棄却される.

第8章

8.1 i) $\hat{a}_1 = 0.0667$, $\hat{a}_2 = 0.2167$

X	0	1	2	3	4
\hat{P}	0.0667	0.2834	0.5000	0.7168	0.9335

iv) $\hat{P}_i < 0$ は $X < -0.3078$ と, $\hat{P}_i > 1$ は $X > 4.3067$ と同値.

8.2 i) $X = -\beta_0/\beta_1$ ii) 最大になる値は $= -\beta_0/\beta_1$, $|dp/dX| = f(0) \cdot \beta_1 = 0.3989\beta_1$ (プロピット), $0.25\beta_1$ (ロジット)

8.3 例 8.5 では大むね成立している. 例 8.2 では, 2.14 倍になっている.

8.4 プロピットに対し 83.3%, ロジットに対し 90.0%.

8.5 i) プロピット(カッコ内は標準偏差)では,

$$Y = -1.4173 + 0.004885 \cdot X_1 - 1.3460 \cdot X_2 - 0.4042 \cdot X_3$$

ロジットでは

$$Y = -2.8062 + 0.009059 \cdot X_1 - 2.3155 \cdot X_2 - 0.5922 \cdot X_3$$

iii) 以下, \hat{A} は近似的(漸近的)に右辺の分布に従うことを表すものとする. プロピットでは, $\hat{P}_i = 0.8474$, $\hat{P}_i - P_i \stackrel{A}{\sim} N(0, 0.06487)$ であるから, 信頼区間は 0.8474 ± 0.4992 . ただし, 上側信頼限界は, 1 を超えるので $(0.3482, 1]$ とする. ロジットでは, $\hat{P}_i = 0.8486$, $\hat{P}_i - P_i \stackrel{A}{\sim} N(0, 0.0715)$ より, 信頼区間は 0.8486 ± 0.5250 . プロピットの場合と同様, 上側信頼

限界は 1 を超えるので (0.3235, 1].

第9章

9.1 i) $P(H_1|S)=9/19=0.47$, $P(H_2|S)=6/19=0.32$, $P(H_3|S)=4/19=0.21$ ii) $(9/19, 6/19, 4/19)$ を事前分布として, $P(H_1|S_1 S_2)=81/133=0.609$, $P(H_2|S_1 S_2)=36/133=0.271$, $P(H_3|S_1 S_2)=16/133=0.120$ iii) $P(S_1 S_2|H_1)=(3/4)^2$, $P(S_1 S_2|H_2)=(1/2)^2$, $P(S_1 S_2|H_3)=(1/3)^2$ から考える.

9.2 $(3/4) \cdot (9/19) + (1/2) \cdot (6/19) + (1/3) \cdot (4/19) = 133/228 = 0.58$, ついでながら, $P(F)$ は $(1/4) \cdot (9/19) + (1/2) \cdot (6/19) + (2/3) \cdot (4/19) = 95/228 = 0.42$.

9.3 $Be(5, 2)$ の平均で $5/7=0.71$.

9.4 精度は事前分布, 尤度についてそれぞれ 2, 4, 事後分布の分散は精度の和の逆数 $1/(2+4)=1/6$. その平均は, 事前分布の平均と $\bar{x}=6.5$ の内分点 $(4 \cdot 5 + 2 \cdot 6.5)/(4+2)=5.5$, $N(5.5, 1/6)$.

9.5 i) $P(D|1, 1, 0)=0.23 \cdot (0.1)(0.7)(1-0.6)/\{0.23 \cdot (0.1)(0.7)(1-0.6)+0.77(0.8)(0.2)(1-0.5)\}=0.095$, $P(\bar{D}|1, 1, 0)=0.905$ ii) $\mathcal{Q}(1, 0, 1)=0.095/0.905=0.105$ 等となる. パターン $(1, 0, 1)$ では, $P(D|1, 0, 1)=0.016$, $P(\bar{D}|1, 0, 1)=0.984$, $\mathcal{Q}(1, 0, 1)=0.016$ 等. 以下, 自ら試みよ. iii) 事前オッズを $\mathcal{Q}=P(D)/P(\bar{D}) (=0.299)$ として, $\mathcal{Q}(S)=\lambda(S) \cdot \mathcal{Q}$. D 対 \bar{D} のオッズは前後で $\lambda=\lambda(S)$ 倍となり, $\lambda \geq 1$ なら症状の観察 S で D の疑いは増す. 例えれば, $\lambda(0, 1, 1)=(1-0.1)(0.7)(0.6)/(1-0.8)(0.2)(0.5)=18.9$ (倍)等. 以下自ら試みよ.

9.6 i) $5(1-w)$, $w+3(1-w)$, $3w+2(1-w)$ を比較すれば, w が $[0, 1/3]$, $[1/3, 2/3]$, $[2/3, 1]$ に属するにしたがい, それぞれ a_3 , a_2 , a_1 をとる. $w=1/2$ なら a_2 . ii) ベイズの定理より, θ_1 の事後確率は z_1 , z_2 , z_3 に応じて, $3/4, 5/11, 3/13$. iii) z_1 のとき a_3 , z_2 のとき a_2 , z_3 のとき a_1 を選択する.

9.7 自ら試みよ.

9.8 アランの確率評価 $(1/2)$ はベイズの定理と整合しない. 整合するには, $P(s|A)=1$ なるべきだが, バーナード, チャールズが釈放される場合, 看守が常に「バーナード」の方の名を挙げると想定するのは不自然. 認知科学的問題.

9.9 i) $Ex(\lambda)$ のモーメント母関数は $\lambda/(\lambda-t)$ ($t < \lambda$). $Ga(a, \lambda)$ のそれは $\{\lambda/(\lambda-t)\}^a$ ($t < \lambda$) ゆえ, 従う. ii), iii) 略(松原(1992)参照).

第10章

10.1 時点 n で原点に戻るために, 上下に等しく k 回ずつ, また左右にも等しく j 回ずつ動き, $n=2k+2j$ をみたさなくてはならない. したがって n が奇数のときは確率 0 である. n が偶数のときは確率 p, q, r, s の 4 項分布と同じであるから, $n=2m$ とおけば

$$\sum_{k=0}^m \frac{(2m)!}{(k!)^2((m-k)!)^2} (pr)^k (qs)^{m-k}$$

となる.

10.2 0 に吸収される確率 $r(a)$ は, 破産問題と同じ考え方により差分方程式 $r(a)=\beta r(a+1)+\gamma r(a)+\alpha r(a-1)$ を満足する. したがって, $r(a)=\beta/(1-\gamma)r(a+1)+\alpha/(1-\gamma)r(a-1)$ を得る. また, 初期条件も破産問題と同じ $r(0)=1$, $r(N)=0$ である.

10.3 i) $p_{jk}=1(j=k=1)=1/3(j \neq 1, j \neq k)=0$ (その他) ii) 愚かなねずみに対し, $4(p=1/4$ の幾何分布の期待値), 普通のねずみに対し, $1/4+\sum_{i=2}^{\infty} i(3/4)(2/3)^{i-2}(1/3)=13/4$, 利口なねずみに対し, $1+2(3/4)(1/3)+3(3/4)(2/3)(1/2)+4(3/4)(2/3)(1/2)=10/4$.

10.4 (10.18) を (10.16) に代入し, 次に (10.16) の両辺に k をかけ $k=1, 2, \dots$ について加えればよい. 微分方程式の解は $m(t)=Ce^{(\mu-\nu)t}$ となり, 初期条件より $C=m(0)=i$ である.

10.5 満たすべき微分方程式は $p_1'(t)=-\lambda p_1(t)$, $p_k'(t)=-k\lambda p_k(t)+(k-1)\lambda p_{k-1}(t)$ となる. あとは与えられた式を直接代入すればよい.

第11章

11.1 自ら試みよ.

11.2 自ら試みよ.

11.3 自ら試みよ.

11.4 (11.16) 式を u_1, u_2 について解くと,

$$x_1^2+x_2^2=-2 \log u_1, \quad u_1=\exp \{-(x_1^2+x_2^2)/2\} \\ x_2/x_1=\tan(2\pi u_2), \quad u_2=\tan^{-1}(x_2/x_1)/2\pi$$

が導かれる. u_1 を x_1, x_2 で偏微分すると, $\partial u_1 / \partial x_1 = -x_1 u_1$, $\partial u_1 / \partial x_2 = -x_2 u_1$ となる. つぎに, 公式 $(\tan^{-1} y)'=1/(1+y^2)$ と合成関数の微分法から, $\partial u_2 / \partial x_1 = -x_2 / 2\pi(x_1^2+x_2^2)$, $\partial u_2 / \partial x_2 = x_1 / 2\pi(x_1^2+x_2^2)$ が得られる. したがって,

$$J=\frac{1}{2\pi} \cdot \frac{1}{x_1^2+x_2^2} \begin{vmatrix} -x_1 u_1 & -x_2 u_1 \\ -x_2 & x_1 \end{vmatrix}=\frac{-1}{2\pi} u_1 \\ =\frac{-1}{2\pi} \exp \left\{ -\frac{1}{2}(x_1^2+x_2^2) \right\}=\frac{-1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2}x_1^2 \right) \times \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2}x_2^2 \right)$$

となる. $f(u_1, u_2)=1(0 \leq u_1 \leq 1, 0 \leq u_2 \leq 1)$ であることに注意すれば, 求める結果が得られる.

11.5 自ら試みよ. 正規性については, 第7章参照.

参考文献

本書を理解するうえで基礎となるもの、その助けとなるもの、さらにくわしく学ぶためのものを中心に参考文献を掲げる。

全面的に基礎となるものは、本シリーズ「基礎統計学」第Ⅰ巻

東京大学教養学部統計学教室(編)『統計学入門』東京大学出版会, 1991.

である。なお、この第Ⅰ巻の参考文献も必要に応じて参照することが有用である。また、関連する内容として、同第Ⅱ巻

『人文・社会科学の統計学』同(近刊)

があり、同一の方法でも分野によって異なった適用法を解説している。

以下、章ごとに、本書の内容程度を中心として、やややさしいもの、やや深い内容のものを挙げよう。なお、ここに挙げたもの以外にも良書があることをお断りしておく。

第1章

森村英典『確率・統計』朝倉書店, 1974.

羽鳥裕久『確率論の基礎』コロナ社, 1974.

脇本和昌『統計学／見方・考え方』日本評論社, 1984.

中村隆英、新家健精、美添泰人、豊田 敬『統計入門』東京大学出版会, 1984.

第2, 4, 5章

竹内 啓『数理統計学』東洋経済新報社, 1963.

廣津千尋『臨床実験データのための統計解析』広川書店, 1992.

廣津千尋『実験データの解析——分散分析を超えて——』共立出版, 1992.

竹内 啓・藤野和建『スポーツの数理科学』共立出版, 1988.

第3章

廣津千尋『分散分析』教育出版, 1976.

奥野忠一、芳賀敏郎『実験計画法』培風館, 1969.

第6章

竹内 啓『確率分布の近似』教育出版, 1975.

清水良一『中心極限定理』教育出版, 1976.

柴田義貞『正規分布』東京大学出版会, 1981.

第7章

ホーエル(村上正康、浅井 晃訳)『初等統計学』培風館, 1981.

参考文献

355

中村隆英、新家健精、美添泰人、豊田 敬『統計入門』東京大学出版会, 1984.

森棟公夫『統計学入門』新世社, 1984.

竹村彰通『現代数理統計学』創文社, 1991.

第8章

コックス(後藤昌司、畠中駿造、田崎武信訳)『二値データの解析』朝倉書店, 1980.

和合 肇・伴 金美『TSPによる経済データの分析』東京大学出版会, 1988.

人文・社会科学への応用では、和書では見出だしにくく、洋書を掲げる。

Amemiya, T., *Advanced Econometrics*, Harvard Univ. Press, 1985.

Maddala, G. S., *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge Univ. Press, 1983.

Pindyck, R., Rubinfeld, D., *Econometric Models and Economic Forecast*, McGrawHill, 1981.

第9章

繁樹算男『ベイズ統計入門』東京大学出版会, 1985.

鈴木雪夫『統計学』朝倉書店, 1987.

松原 望『統計的決定』放送大学教育振興会, 1992.

松原 望『入門ベイズ統計』東京図書, 2008.

松原 望『ベイズ統計学概説』培風館, 2010.

渡部 洋『ベイズ統計学入門』福村出版, 1999.

第10章

森村英典『確率・統計』朝倉書店, 1974.

脇本和昌『統計学／見方・考え方』日本評論社, 1984.

Feller, W., *An Introduction to Probability Theory and its Applications*, Wiley
フェラー(河田龍夫監訳)『確率論とその応用』I, II, 紀伊國屋書店, 1960,
1961.

Karlin, S., *A First Course in Stochastic Processes*, Academic Press

カーリン(佐藤健一・佐藤由身子訳)『確率過程講義』産業図書, 1974.

第11章

クヌース(渋谷政昭訳)『準数値算法／乱数』サイエンス社, 1981.

伏見正則『乱数』東京大学出版会, 1989.

[各章における関連文献] 内容に関連して引用された文献は次の通りである。

第2, 4, 5章

- 廣津千尋『離散データ解析』教育出版, 1982.
- 廣津千尋『統計的データ解析／工学、医学、薬学、社会データの実例による説明』日本規格協会, 1983.
- 廣津千尋『実験データの解析』共立出版, 1992.
- 廣津千尋(編)『統計的方法』日本規格協会, 1993.

第6章

- 竹内 啓, 大橋靖雄『統計的推測——2標本問題』日本評論社, 1981.
- 水内知子, 水内 厚, 田口隆彦, 木田厚瑞「幼若ラットにおける肺切除後の代償発育：葉間差について」, 日本胸部疾患学会雑誌 29巻増刊号, 1991.

第9章

- Chernoff H., Moses L., *Elementary Decision Theory*, John Wiley & Sons
- チャーノフ, モーゼス(宮沢光一訳)『決定理論入門』, 2001.
- 松原 望『意志決定の基礎』(新訂版)浅倉書店, 1985.
- D. Berry 'Bayesian Methods in Phase III Trials' *Drug Information Journal*, Vol. 25, pp. 345-368, 1991.

第10章

- 刈屋武昭, 佃 良彦, 丸 淳子編著『日本の株価変動』東洋経済新報社, 1989.

- 読者の便宜のため、第I巻で紹介した事典、解説書、手引き、数表を挙げておく。
- 池田 央(編)『統計ガイドブック』新曜社, 1989.
 - 上田尚一『統計用語辞典』東洋経済新報社, 1981.
 - 竹内 啓(編)『統計学辞典』東洋経済新報社, 1988.
 - 統計教育推進会『統計小辞典』日本評論社, 1989.
 - 中山伊知郎(編)『現代統計学大辞典』東洋経済新報社, 1963.
 - 溝口敏行(編)『統計ガイドブック』日本経営協会, 1989.
 - 廣松 毅(他)『ロータス1-2-3による統計入門』朝倉書店, 1988.
 - 和合 敏・伴 金美『TSPによる経済データの分析』東京大学出版会, 1988.
 - 大橋靖雄・市川伸一『SASによるデータ解析入門』東京大学出版会, 1987.
 - 三宅一郎・山本嘉一郎『SPSS統計パッケージ』(I, II)東洋経済新報社, 1976.
 - 統計数値表編集委員会『簡約統計数値表』日本規格協会, 1977.

索引

重要項目は大項目とし、派生項目を細分項目として整理して掲げた。たとえば、「指數分布の期待値」は「期待値」で参照できる。

なお、⇒は、関連項目への参照、「…」見よは、同義語への参照、「…」も見よは、他の重要関連項目への参照を示す。

頁数については、ffはその頁以降を意味する。

ア

一般化逆行列	42	見よ。
一般線形モデル	111 ff, 115	オッカムの剃刀 205
一般平均	87, 97, 100	オッズ 173, 237
移民	297	事後— 274
ANOVA	84, 89, 92	対数— 237
アーラン分布	299	力
一元配置分散分析	84 ff, 86, 88, 102, 108	回帰係数 62
一元配置(のモデル)	31, 33, 36, 39, 47, 49, 59, 63, 65, 75	回帰分析 31, 34, 50, 76, 78 —の2標本問題 77
一次従属	68	原点を通る— 78
一次変換	67	回帰モデル 76
1標本問題	183	2次— 59
位置母数	204	χ^2 (カイ二乗)成分 158
ロジスティック分布の—	162	群間の— 158
一様最強力検定 ⇒ 最強力検定		群内の— 158
一様最強力不偏検定	200	χ^2 適合度検定 「適合度検定」 を見よ。
一様性	59, 94, 149, 150	χ^2 適合度統計量 「適合度統計量」を見よ。
—の帰無仮説	157	χ^2 統計量 「適合度統計量」 を見よ。
—の検定	146	
多项分布における—	170	
一様分布	18, 22	
—の限界の最尤推定	143	
ペイズ統計学における—	255, 267	
乱数発生における—	309, 312, 318	
一様乱数	146, 311 ff, 320, 331	
一様乱数表	309	
一对比較表	170	
一致(性)	128	
一致推定量	128, 240, 246	
OC曲線 「作用特性曲線」を		—の推定 243

- 原因の—— 251, 254
 事後—— 251, 252 ff
 事前—— 251 ff
 確率化 86
 確率過程 277 ff
 正規—— 286
 出生死滅過程 278, 293, 296
 ff
 ブラウン運動 278, 283 ff,
 286 ff
 ポアソン過程 278, 293 ff
 マルコフ連鎖 286 ff
 ランダム・ウォーク 278 ff
 待ち時間の—— 298 ff
 確率関数 2
 確率行列 288
 確率収束 128-130, 137
 確率の推定 241, 243
 ——値の漸近分布 243, 246
 二項的な場合の—— 241
 確率分布(各論)
 アーラン分布 299
 一様分布 「一様分布」を見よ.
 F 分布 61, 92
 χ^2 分布 18, 326
 ガンマ分布 11, 17, 23, 257,
 275, 326 ff
 幾何分布 4 ff, 17, 22, 142
 コーシー分布 208, 228
 指数分布 「指数分布」を見よ.
 自然な共役分布 255 ff
 正規分布 「正規分布」を見よ.
 多項分布 141, 148, 150, 171
 超幾何分布 4, 7, 9, 22
 t 分布 208
 統計的決定理論における—— 266 ff
 二項分布 「二項分布」を見よ.
 非心 F 分布 62, 64
 非心 t 分布 64, 188
 標準正規分布 「標準正規分
- 布」を見よ.
 負の二項分布 「負の二項分布」を見よ.
 ベータ分布 255 ff
 ポアソン過程の—— 294
 ポアソン分布 「ポアソン分布」を見よ.
 マルコフ連鎖における—— 291
 ロジスティック分布 112 ff,
 161, 253 ff, 323
 ワイブル分布 323
 ランダム・ウォーク 278 ff
 待ち時間の—— 298 ff
 確率密度関数 2
 過誤 178
 第1種の—— 179, 180, 181,
 194, 214, 264
 第2種の—— 179, 180, 181,
 193, 194, 264
 分類の—— 274
 片側(右, 左)検定 60, 183 ⇒ 検定
 片側対立仮説 ⇒ 検定
 傾き 77
 カテゴリ(カテゴリカル) 171
 株価 301 ff
 加法性 95, 99
 仮平均 90, 101
 カルマン・フィルター 262
 間隔尺度 171
 完全無作為化法 86, 96
 観測度数 145, 148
 官能検査 161 ff
 ガンマ関数 17
 ガンマ分布 11, 17, 23
 ——の形状パラメータ 326
 ——の尺度パラメータ 326
 ウィルソン・ヒルファーティの
 近似 327
 指数分布と—— 23
 ベイズ統計学における—— 257, 275
 乱数発生における—— 326
 ガンマ乱数 326 ff
- 規格化 253
 繼承分布 4, 6 ff, 17, 142
 ——の期待値 22
 ——の最尤推定 142
 ——のフィッシャー情報量
 142
 離却 178 ff, 265
 ——域 60, 179
 ——点 179, 182
 模似乱数 173, 311
 規準化変数 61
 希少性 295
 確率変数 1, 253, 277, 309
 期待損失 259, 260
 ——の和 3, 23
 最小事後—— 260, 266, 268
 損失関数と—— ⇒ 損失関数
 期待値 3, 13, 15, 19, 31, 102
 一様分布の—— 22
 幾何分布の—— 22
 指数分布の—— 23
 二項分布の—— 5, 22
 負の二項分布の—— 22
 平方和の—— 109
 ベータ分布の—— 257
 ポアソン分布の—— 22
 モーメント母関数による—— 19
 ランダム・ウォークの—— 283
 倍無仮説 60, 138, 139, 140,
 145, 148, 165, 178 ff, 179
 逆関数法 321
 級間平方和 89, 90
 吸收壁 280
 共役分布 「自然な共役分布」
 を見よ.
 共分散 3
 分散 —— ⇒ 分散
 局外母数 195
 極座標法 「ボックス・ミュラ
 ー法」を見よ.
 区間推定 129
 「信頼区間」も見よ.
 一元配置の処理効果の—— 93

- 正規分布の—— 205
 二元配置の処理効果の—— 104
 母平均に関する—— 205
 母平均の差の—— 82
 母分散の比の—— 84
 クラメール(G.) 123
 クラメールの公式 64
 クラメール(H.) 123
 クラメール・ラオの下限 123
 ff, 131, 133
 クラメール・ラオの不等式 126
 最尤推定量の—— 131
 繰り返し数 88
 繰り返し測定(のモデル) 26,
 32, 36, 39, 43, 50, 53, 58, 59,
 62, 75, 116, 126, 127, 204
 クロス・バリデーション法 66
 群間 64, 158
 ——の x^* 成分 158
 群内 64, 158
 ——の x^* 成分 158
 計画行列 32, 69, 75
 形状パラメータ 18, 326
 t 標本問題 84, 184
 ——としての分散分析 84
 計量分類 266, 272, 274
 結果 252
 原因 252
 検出力(関数) 64, 136, 177 ff,
 185 ff, 196, 215, 218
 最強力検定の—— 196
 ノンパラメトリック検定の—— 218
 標本の大きさと—— 195 ff
 検出力関数(各論)
 f_1 下での—— 215
 正規分布の分散に関する 2 標本
 問題の—— 191
 正規分布の母平均に関する検定
 の—— 185, 190
 正規分布の母平均に関する t 検
 定の—— 187
 Z にもとづく 2 標本検定の一
- 188, 200
 t 検定の—— 187, 215
 二項分布の—— 192, 193,
 200
 2 標本問題の—— 193
 2 標本 t 検定の—— 188
 検定(一般論) 129, 264
 片側—— ⇒ 検定(各論)
 兩側—— ⇒ 検定(各論)
 クラメール・ラオの下限 123
 ——統計量 179
 —における過誤 178 ff
 —の検出力 ⇒ 検出力
 —の漸近論 135 ff
 —方式 178
 最強力—— 196 ff
 統計的決定問題としての—— 264
 検定(各論) 「2 標本検定」も
 見よ.
 一様性の—— 146, 156 ff
 ウィルコクソン—— 221
 ウエルチの—— 82
 順位和—— 221 ff, 229
 正規性の—— 223 ff
 t 標本問題 84, 184
 ——としての分散分析 84
 計量分類 266, 272, 274
 結果 252
 原因 252
 検出力(関数) 64, 136, 177 ff,
 185 ff, 196, 215, 218
 最強力検定の—— 196
 ノンパラメトリック —— 218 ff, 218, 222
 パートレットの—— 94
 符号 —— 218, 229
 プロピット・モデルの—— 249
 分割表の—— 162
 分散分析 —— 89, 91
 ポアソン分布に対する最尤法に
 よる—— 136
 マン・ホイットニーの—— 199
 Z にもとづく 2 標本検定の一
- ロジット・モデルの—— 249
 ケンドール過程 297
 効果 87
 交互作用 94 ff, 99 ff, 103 ff,
 109, 159, 172 ff
 公平 281
 誤差 25 ff, 30 ff, 41, 50, 53,
 100, 206, 208, 211 ff
 ——項 234
 —分散 51 ff
 —平方和 89 ff
 —変動 89
 偶然—— 25, 30, 53
 サンプリング—— 33
 実験—— 100, 104 ff
 コーシー分布 29, 207 ff
 ——の平均不存在 208
 最尤推定量 228
 標本中央値 228
 フィッシャー情報量 228
 亂数発生における—— 323
 固定標本数 262
 固有値 56 ff, 69, 72
 固有ベクトル 56 ff, 72
 固有方程式 58, 73
 固有値問題 72
 コルモゴロフ, A 125
 サ
- 最強力検定 196 ff
 一様—— 197, 200
 正規分布に対する一様—— 200
 減近的に一様—— 136
 二項分布に対する一様—— 199
 最小事後期待損失 260, 266,
 268
 最小二乗解 40 ff, 121
 最小二乗推定量 55, 74
 ——に関するガウス・マルコフ
 の定理 40
 ——の分散 49
 最小二乗法 29, 32, 36 ff, 40,

- 47, 61, 75, 117
 最小値 18
 最小分散 55, 56, 59, 206
 最小分散不偏推定量 43, 126, 204
 最大値 18
 採択 178 ff, 180, 265
 最適 265
 最適水準 102
 　　—組合せ 96
 最尤推定量(一般論)
 　　—の一致性 128
 　　—のクラメール・ラオの下限 131
 　　—の最適性 128 ff
 　　—の漸近分布 131, 135
 　　—の漸近有効性 128
 　　多次元母数の— 141
 尤度の最大値としての—
 　　「尤度(関数)」を見よ。
 最尤推定量(量, 値) 116, 118 ff, 121, 133, 135, 137, 159, 164, 208 ff, 228, 246
 一様分布の限界の— 143
 幾何分布に対する— 142
 指数分布に対する— 143
 正規分布に対する— 134
 渐近分布に対する— 135, 136
 多項分布に対する— 141, 148
 二重分類に対する— 140
 標準偏差の— 143
 ブラッドリー・テリーのモデルの— 168
 プロビット・モデルの— 238
 分割表に対する— 141
 ポアソン分布に対する— 132, 136
 有効推定量としての— 133
 ロジット・モデルの— 298
 最尤法(最尤推定に同じ) 115
- ff, 211, 237
 最良線形不偏推定量 29 ff
 最良不偏推定量 117
 SAS 241
 差分方程式 304
 作用特性曲線 182
 3元配置 96
 残差 51, 76
 　　—平方和 52, 53, 63, 65, 72, 88
 3囚人問題 274
 3重分類表 170
 算術乱数 311
 三すくみ 167, 170
 サンプリング誤差 33
 サンプル・バス 「標本関数」
 　　を見よ。
 試行 4
 事後確率 252, 274
 　　—分布 253 ff, 258 ff
 自己相関係数 ⇒ 「相関係数」
 指数分布 11, 16 ff, 23, 300, 320
 　　—とガンマ分布 23
 　　—に対する中心極限定理 23
 　　—の最尤推定 143
 　　—のモーメント母関数 23
 サービス時間の— 300
 ベイズ統計学における—
 　　275
 待ち行列の— 300
 亂数発生における— 323
 指数乱数 323
 事前確率 251
 　　—分布 253 ff, 259, 273
 自然な共役分布 273
 　　—ガンマ分布の— 275
 　　—正規分布の— 273
 　　—としてのガンマ分布 275
 　　—としての正規分布 257
 　　—としてのベータ分布 255 ff
 実験 96
- 実験計画法 87
 実験誤差 100, 104
 質的データ 232
 Cp 基準 66
 四分位点 202
 四分位範囲 212
 死亡率 298
 射影子 43, 53, 61
 尺度
 　　—名義— 171
 　　—順序— 171
 尺度パラメータ 18, 326
 尺度母数 162
 三周期 312, 313, 316, 318
 周期的(性) 312
 修正項 90, 105
 集団遺伝学 289, 292
 自由度 18, 52, 56, 82, 91 ff, 155
 十分統計量 258
 ベイズ統計学における—
 　　258
 周辺対称性 ⇒ 対称性
 主効果 94, 99, 102, 104, 109, 172
 出生死滅過程 278, 296 ff, 297, 299
 —における移民 297
 ケンドール過程 297
 純出生過程 297
 フェラー・アレイ過程 297
 ユール過程 297
 出生率 298
 受容 178 ff
 願位(ランク)とも 61, 67 ff, 219
 —相関係数 ⇒ 相関係数
 —和 220
 —和検定 221 ff, 229
 併合 219
 純出生過程 297, 306
 順序統計量 210, 211, 227 ff
 単対称性 ⇒ 対称性
 小数の法則 「ポアソンの小数の法則」を見よ。

- 乗算型合同法 313, 331
 乗数(合同法の) 313
 状態 267, 286
 統計的決定理論における—
 　　267
 マルコフ連鎖における—
 　　286
 情報行列 「フィッシャーの情報行列」を見よ。
 消費者危険 182
 乗法モデル 115
 情報量 「フィッシャー情報量」
 　　を見よ。
 情報量基準 66
 処理 31, 85, 100, 161
 　　—効果 93
 　　—群 80
 　　—平均 104, 109
 信頼区間 82, 84, 93, 108, 109, 205
 　　「区间推定」も見よ。
 信頼係数 82, 84, 93, 103
 信頼限界 93, 103, 109
 信頼性 275
 シンプソンのバラドックス 172, 173
 推移確率(行列) 288 ff, 290
 　　—に対するチャップマン・コルモゴロフの方程式 290
 水準 31, 34, 47, 85, 97, 99, 102, 107
 　　—組合せ 94 ff, 171
 　　—平均 93
 　　因子の— 172
 推定
 　　—可能 38 ff
 　　—関数 39, 40, 48, 49
 　　—空間 42, 45, 49
 信頼区間にによる— ⇒ 区間推定
 点— 「点推定」を見よ。
 統計的決定問題としての—
 　　264
 ベイズ— 259 ff
 　　—モーメントによる— 176
 生産者危険 182
- 推定量 122, 123
 一致性 ⇒ 一致推定量
 最尤推定 ⇒ 最尤推定量
 渐近的有効性 128, 133
 不偏性 ⇒ 不偏推定量
 フーバー型— 212
 ホッジ・レーマン— 28
 ステム・アンド・リーフ 「幹葉表示」を見よ。
 節減(約)の原理 205
 絶対損失 260, 261, 262
 切片 77
 セル 113, 141
 游近(的) 135, 137
 　　—分散 137, 139
 　　—に有効 131
 　　—有効性 129
 游近分布 128, 131, 133, 146, 150, 154, 240, 246
 二項分布に対する— 154, 192
 標本中央値の— 210
 ロジット、プロビット・モデルの— 246
 ロジット、プロビット・モデルによる推定確率の— 243, 246
 線形 25, 29, 31, 32
 　　—仮説 59, 60
 　　—最小分散不偏推定量 42
 　　—推測論 32
 　　—制約 59
 線形確率モデル 249
 線形合同法 311 ff, 315, 316
 線形判別関数 269, 272, 274
 線形モデル 25 ff, 31, 32, 59, 65, 111, 170
 　　—ガウス・マルコスの定理 40 ff
 　　—の計画行列 32
 　　—の最小二乗推定 36 ff
 　　—の正規方程式 37 ff
 一元配置の— 「一元配置のモデル」を見よ。
 繰り返し測定の— 「繰り返し測定のモデル」を見よ。

対数—— 170
 多項式の—— 「多項式のモデル」を見よ。
 単回帰の—— 「単回帰のモデル」を見よ。
 尖度 223 ff, 229
 標本—— 224 ff
 母—— 224
 像 68, 70
 相関係数
 自己—— 301, 303
 順位—— 216
 積率—— 216
 総平方和 89, 100
 総変動 89
 増分 286
 測定誤差 33, 204
 素数 313 ff
 ——判定法 314
 メルセンヌ—— 314
 乱数発生のための—— 313
 損失(関数) 259, 260 ff, 264
 絶対—— 260, 262
 単純—— 260, 262
 統計的決定理論における—— 264
 ベイズ判別における—— 270
 平方—— 260, 262
 非対称絶対—— 260
 期待—— 259 ff
 タ
 退化 68
 対称 43
 対照群 80
 対照実験 79, 80
 対称性 52, 69, 162 ff, 229
 周辺—— 165 ff
 庫—— 165 ff, 170
 対数オッズ 237
 対数線形モデル 115, 170 ff, 173
 大数の法則 28, 129, 130, 214
 対数尤度 117

——とフィッシャー情報量 122
 エフィシェント・スコアを導く—— 138
 プロビット、ロジット・モデルの—— 247
 ポアソン分布の—— 132
 尤度比検定を導く—— 136
 対立仮説 81, 178 ff, 189, 196, 199 ff
 片側対立仮説 \Rightarrow 検定
 兩側対立仮説 \Rightarrow 検定
 多因子実験 94, 107
 多元配置 96
 多項式モデル 31, 34, 36, 37, 39
 多項選択 231
 多項反応 231
 多項分布(のモデル) 115, 118, 141, 150, 171, 241
 ——に対する最尤推定(量) 141, 148
 定常(性) 4, 294
 定常分布(マルコフ連鎖の) 291, 298, 300
 に対する尤度比検定 148
 多次元 139, 140
 —母数 139
 多次元(多変量)正規分布 74, 270
 多次元疎結晶構造 316, 318, 328
 多重マルコフ連鎖 288
 たたみこみ 275
 ダミー変数 232
 単一因子 96
 ——実験 107
 単回帰(分析) 51, 76
 単回帰(のモデル) 44, 45, 50, 53, 62, 76, 111, 121
 簡単損失 260, 262
 チェビシェフ, P 125
 チェビシェフの不等式 28
 チャップマン・コルモゴロフの方程式 290
 中央値 28, 202
 標本—— 208 ff, 217, 228
 母—— 210

中心極限定理 12, 20, 23, 30, 130, 192, 214, 240, 286, 328, 332
 指数分布に対する—— 23
 超幾何分布 4, 7, 9, 22
 —に対する二項近似 22
 直交表 107
 2水準—— 107
 直交補空間 52, 70 ff
 TSP 241
 t検定 60, 62, 213
 —の検出力関数 187
 —のための t 分布 60
 —のための t 分布表 336
 t統計量による—— 62, 64, 187
 定常(性) 4, 294
 定常分布(マルコフ連鎖の) 291, 298, 300
 t統計量 62, 64, 187
 t分布 60, 205, 208, 327
 —表 337
 t統計量による—— 62, 64, 187
 亂数発生における—— 327
 非心—分布 64, 188
 適合度検定(一般論)
 エフィシェント・スコアによる
 — \Rightarrow エフィシェント・スコア」を見よ。
 χ^2 適合度統計量による—— \Rightarrow (χ^2)適合度検定」を見よ。
 尤度比による—— \Rightarrow 尤度比検定」を見よ。
 尤度比による—— \Rightarrow 尤度比検定」を見よ。
 (x²)適合度検定 120, 145 ff, 147, 149, 172-175
 一様性の—— 157
 完全独立の—— 154
 交互作用を入れた独立性の—— 159, 160
 対称性の—— 164, 165
 対数線形モデルの—— 172
 多項分布の—— 150
 独立性の—— 151 ff

負の二項分布の—— 176
 ポアソン分布の—— 149, 150, 175
 (χ^2)適合度統計量 141, 145, 146, 152
 —の分解 157
 対称性の—— 164, 166
 独立性の—— 152 ff
 ブラッドリー・テリーのモデルの—— 168
 点推定 206 ff, 252 ff, 257, 261, 265, 267, 269 ff, 274
 統計的決定 251 ff
 —問題 260, 266, 270
 統計的推測 264
 同時伯頃域 263
 等分散(性) 28, 35, 41, 108
 —仮説 94
 特異値分解「スペクトル分解」を見よ。
 独立性(確率論, 確率分布論の) 4, 56
 独立(モデル) 115, 141, 142, 148, 151, 154, 164, 170
 —性の χ^2 適合度検定 151, 164
 二重分類 140, 148, 151
 二重百換 83
 二値性 4, 295
 完全—— 159
 交互作用を入れた—— 159, 160
 分割表の—— 141, 142
 独立増分 286
 ドット記法 88
 度数
 観測—— 145 ff
 理論—— 145 ff
 トリム平均 211, 212, 213, 229
 トレース 69
 ナ
 2元配置 65, 96, 97, 99, 104, 105, 106, 109
 —分散分析 96 ff
 —繰り返しのない—— 104 ff
 二項係数 5
 二項定理 5
 二項的 232
 二項反応 231
 二項分布(モデル) 4, 9, 20, 22, 111, 118, 152, 154, 178, 185, 217, 294
 —に対する一様最強力検定 199
 —に対する検定 178 ff, 192 ff, 200
 —に対する漸近分布 154
 —の期待値 5
 —の正規近似 21
 —の比較 153, 155
 —の分散 5
 —のペイズ推定 216
 符号検定における—— 217
 ベイズ統計学における—— 255, 261
 ランダム・ウォークにおける—— 279
 2次回帰モデル 59
 2次多項式(モデル) 35, 38, 47, 75
 二重分類 140, 148, 151
 二重百換 83
 二値データ 231 ff, 232, 236
 2標本問題(検定) 79 ff, 80, 86, 108, 183, 184, 188
 分割表の—— 141, 142
 独立増分 286
 ドット記法 88
 度数
 観測—— 145 ff
 理論—— 145 ff
 トリム平均 211, 212, 213, 229
 トレース 69
 ナ
 2元配置 65, 96, 97, 99, 104, 105, 106, 109
 —分散分析 96 ff
 —繰り返しのない—— 104 ff
 二項係数 5
 ネイマン・ピアソンの検題 135, 197
 抽取検査 178, 182
 ノイマン(フォン・) 330
 ノー・データ問題 268, 274
 ノンパラメトリック検定 217 ff, 222
 符号検定 229
 順位和検定 229
 ノンパラメトリック(的) 216, 222
 ハ
 排他的論理和 317
 配列 317
 破産問題 279 ff, 288
 —の破産確率 282
 外れ値 202, 203
 パートレットの検定 86, 94
 パラメータ 10
 パラメトリック(的) 25, 222
 反射壁 289, 293
 半数致死量 231, 244
 判別 268, 269
 判別 \Rightarrow ベイズ判別
 ピアソン, E. 196
 ピアソン, K. 196
 比較 79
 2母集団の—— \Rightarrow 2標本問題
 非心度 64
 非心 F 分布 62, 64
 非心 t 分布 64, 188
 非対称性 223
 非復元抽出 7
 ピュッフォンの針 330
 標準化 15
 標準誤差 76
 標準正規分布 14 ff, 60, 81, 112, 154, 173, 185, 192, 227, 235 ff
 分散に関する—— 191
 ニュートン, I. 239
 ニュートン法 238, 239
 ネイマン, J. 197
 ネイマン・ピアソンの検題 135, 197
 ノイマン(フォン・) 330
 ノー・データ問題 268, 274
 ノンパラメトリック検定 217 ff, 222
 符号検定 229
 順位和検定 229
 ノンパラメトリック(的) 216, 222
 ハ
 排他的論理和 317
 配列 317
 破産問題 279 ff, 288
 —の破産確率 282
 外れ値 202, 203
 パートレットの検定 86, 94
 パラメータ 10
 パラメトリック(的) 25, 222
 反射壁 289, 293
 半数致死量 231, 244
 判別 268, 269
 判別 \Rightarrow ベイズ判別
 ピアソン, E. 196
 ピアソン, K. 196
 比較 79
 2母集団の—— \Rightarrow 2標本問題
 非心度 64
 非心 F 分布 62, 64
 非心 t 分布 64, 188
 非対称性 223
 非復元抽出 7
 ピュッフォンの針 330
 標準化 15
 標準誤差 76
 標準正規分布 14 ff, 60, 81, 112, 154, 173, 185, 192, 227, 235 ff
 分散に関する—— 191
 ニュートン, I. 239
 ニュートン法 238, 239
 ネイマン, J. 197

- 標本中央値 211
 標本の大きさ 193 ff, 194, 200
 正規分布の母平均の検定 196
 二項分布の μ の検定 196, 200
 標本分布
 残差平方和の—— 56
 正規線形モデルと—— 55
 ff
 「漸近分布」も見よ。
 標本平均 204
 標本歪度 224 ff
 品質管理 182
 フォン・ノイマン, J. 330
 フィッシャー, R. 83
 フィッシャー情報量 120 ff
 エフィシェント・スコアによる検定と—— 141
 幾何分布の—— 142
 クラメール・ラオの下限と—— 126
 繰り返し測定に対する—— 127
 コーナー分布の—— 228
 対数尤度と—— 122
 ポアソン分布の—— 127
 フィッシャーの情報行列 123, 133, 134
 正規分布の—— 134
 フェラー・アレイ過程 297, 300, 305
 フェルミ, E. 322
 不確実性下の意思決定 265
 不均一性 \Rightarrow 一様性
 復元抽出 9
 符号検定 218, 229
 負の二項分布 4, 6 ff
 ——に対する χ^2 適合度検定 176
 ——の期待値 22
 ——の推定 176
 フーバー型推定量 229
 部分空間 61
 不偏性(検定の) 28, 123, 200
 不偏検定 200
 一様最強力—— 200
 不過推定量 29, 56 ff, 123, 204
 ff, 208
 回帰直線の誤差分散の——
 最小分散—— 43
 最良線形—— 29
 線形モデルの誤差分散の——
 52, 53, 56, 58
 線形モデルに対する—— 40
 分散分析における誤差分散の——
 —— 94
 不偏性(推定の) 28, 29, 35, 240
 不偏分散 53, 59, 94
 不変分布(マルコフ連鎖の) 291
 ブラウン運動 278, 283 ff, 286
 ff, 301
 ——仮説 301 ff
 ——の期待値 283
 ——の独立増分性 286
 ——の分散 283
 フラクショナル—— 287
 フラクタル 287
 プラセボ 183, 253, 261, 266
 ブラッドリー・テリーのモデル 163, 167, 170, 176
 ——の最尤推定 167
 ——の適合度検定 176
 ——の適合度統計量 167
 ブル・ランク 40, 43, 49, 55, 68
 不連続補正 328
 ブロック 85, 86
 プロビット(モデル) 113, 114, 232 ff, 236
 ——による確率の推定 241, 245
 ——による推定結果 242
 ——の検定 249
 ——の最尤推定 238, 245
 線形モデルの導入としての—— 113 ff
 分割表 118, 141, 148, 151, 161, 170
 ——の独立性 141, 142
 3次元—— 170 ff
 正方—— 170
 分散(確率変数, 確率分布の) 3
 一様分布の—— 22
 χ^2 分布の—— 59
 正規分布の—— 13
 二項分布の—— 5, 22
 標準正規分布の—— 15
 ベータ分布の—— 257
 ポアソン分布の—— 22
 モーメント母関数による—— 22
 ランダム・ウォークの—— 283
 分散(線形モデルの) 35
 ——共分散行列 36, 52, 272
 ——の一様性 94
 誤差—— 51 ff, 59, 77, 86, 94
 最小—— 28, 55, 56, 59, 206
 最小二乗推定量の—— 49
 ff
 最尤推定量の漸近—— 137
 不偏推定量の—— 123
 分散比 64, 83, 84
 分散分析(法) 31, 34, 64, 84, 87, 88 ff, 172
 分散比 64, 83, 84
 分布関数 2
 分布によらない 216
 分類 268
 ——過誤 274
 計量—— 269
 分類表 「分割表」を見よ。
 平均の検定 137
 平均平方 91
 平均ベクトル 36
 併合順位 219

- 併合推定量 81
 ベイジアン 「ベイズ(的)」を見よ。
 ベイズ, T. 254
 ベイズ(的) 251
 ——決定 251
 ——決定方式 268
 ——更新 257, 262, 273
 ——診断 273
 ベイズ推定 258 ff
 各損失関数による—— \Rightarrow 損失関数
 正規分布の—— 261, 262
 二項分布の—— 261
 ベイズの定理 251 ff, 254, 257, 267, 273 ff
 ベイズ判別 268 ff, 270
 平方損失 260, 262
 平方和 90 ff, 99 ff
 ——の期待値 109
 級間—— 89 ff
 群内—— 63
 交互作用—— 101
 誤差—— 89 ff
 残差—— 52 ff, 56
 総—— 89
 平方和の分解 90, 99, 101 ff, 105, 109
 平均 84
 母集団 183
 ——平均 84
 母数(「パラメータ」とほぼ同義) 10, 36, 86, 100, 122, 133, 139, 177, 183
 積分—— 273
 多次元の—— 133, 139
 母分散 81
 ——についての検定 \Rightarrow 検定
 ——についての検定の検出力 191
 母平均 80
 ——の差の推定 82
 ——の差の平均 80
 ベータ関数 18, 255
 ベータ分布 18
 ベイズ統計学における—— 255, 257, 261, 273
 別名法 324 ff
 ベルタランフィー, L. 248
 ベルヌーイ試行 4, 6, 7, 9, 11, 20, 279, 294
 ——とポアソン過程 294
 偏差値 15
 偏差二乗和· 36, 37
 變動 「平方和」を見よ。
 變動 63, 89, 100
 ポアソン過程 278, 293 ff, 298
 ポアソンの少数の法則 9
 ポアソン分布 4, 9 ff, 16, 19, 219
 マハラノビス 125
 22, 127, 131 ff, 136 ff, 149, 294
 —に対するエフィシェント·スコアによる検定 139
 —に対する最尤推定 132
 —に対する適合度検定 150, 175
 —の期待値 19
 —の分散 19
 —のモーメント母関数 19
 その独立確率変数の和 23
 亂数発生における—— 327
 法(合同法の) 312
 法則収束 246
 母集団 183
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ, A. 125
 マルコフ性 288, 293
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30
 指数分布の—— 23
 正規分布の—— 19
 独立確率変数の—— 20
 ポアソン分布の—— 19
 和の—— 20
 モンテ・カルロ法 322, 330
 マハラノビス 125
 マルコフ連鎖 278, 286 ff, 288
 ff
 ——に対するチャップマン・コルモゴロフの方程式 290
 ——の状態 286
 ——の推定確率(行列) 288
 ——の定常分布 291
 ——の不变分布 291
 齊時—— 288
 多直—— 288
 マン・ホイットニー検定 「頃位和検定」を見よ。
 幹葉表示 201, 203, 229
 密度関数 2, 122
 未定係数法 119
 無作為化 86
 無相関(性) 28, 35, 41, 52, 57
 名義尺度 171
 メディアン 260, 262
 メンデル, G. 132
 メンデルの法則 132
 モジュロ(mod) 311
 モード 260 ff
 モーメント 223
 平均の周りの—— 223
 モーメント法 176
 モーメント母関数 15, 19 ff, 23
 ——による期待値の導出 19
 モーメントの—— 30

217	フル・―― 40, 43, 49, 55, 68
—性検定 181	乱数 202, 307 ff
—点 179	一様 311 ff
有効推定量 126 ff, 133, 150	χ^2 327
—としての最尤推定量 133	ガンマ 326 ff
尤度(関数) 115, 121, 131, 132, 138	擬似 311
—の最大化と最尤推定量 「最尤推定量」を見よ。	コーシー分布に従う 323
—の対数 「対数尤度」を見よ。	算術 311
ベイズ統計学における 253	指數 323 ff
尤度比 136, 274	正規 319
—の対数 138	t 分布に従う 327
エフィシェント・スコアの導出 138	標準正規分布に従う 323
尤度比検定 120, 136 ff, 140, 141, 143, 146, 149	ロジスティック分布に従う 323
一様性の 149	ワイル分布に従う 323
セル確率の 149	乱数サイ 309
対数線形モデルの 172	乱数シミュレーション 322
適合度検定としての 120	乱数表(付表) 308
尤度方程式 117, 129, 132	ランダマイズ 86
ユール過程 \Rightarrow 純出生過程	ランダム 147
要因実験 94, 96	ランダム・ウォーク 278 ff, 288
予測不能性 309	—と破産問題 279 ff
予測分布 273	静止のある 305
ラ	2 次元の 279, 305
ラオ 125	ランダム・サンプル 33
ラプラス 254	離散(時間) 277
乱塊法 86	離散型 1
ランク 「順位」を見よ。	離散分布 4 ff, 19
ランク(階数) 38	乱数発生における 324
—落ち 39, 42, 47, 48, 68	リモート・センシング 272
	理由不十分の原則 252
	両側検定 \Rightarrow 検定(各論)
	両側対立仮説 \Rightarrow 検定(各論)
	量・反応関係 231, 233, 241, 243

本書の執筆者

矢島美寛 東京大学大学院経済学研究科教授

廣津千尋 明星大学理工学部教授

藤野和建 前東京大学教養学部教授(1992年6月逝去)

竹村彰通 東京大学大学院情報理工学研究科教授

竹内啓 東京大学名譽教授

繩田和満 東京大学大学院工学系研究科教授

松原望 聖学院大学大学院政治政策学研究科、東京大学名譽教授

伏見正則 南山大学情報理工学部教授

自然科学の統計学

基礎統計学 III

1992年8月18日 初版

2010年2月26日 第13刷

[検印廃止]

編者 東京大学教養学部統計学教室

発行所 財団法人 東京大学出版会

代表者 長谷川寿一

113-8654 東京都文京区本郷7-3-1 東大構内

電話 03-3811-8814・振替 00160-6-59964

印刷所 株式会社三秀舎

製本所 矢嶋製本株式会社

© 1992 Statistics Section, Department of Social Sciences, College of Arts and Sciences,
The University of Tokyo
ISBN 978-4-13-042067-9 Printed in Japan

〔日本複写権センター委託出版物〕

本書の全部または一部を無断で複写複製(コピー)することは、
著作権法上の例外を除き、禁じられています。本書からの複写
を希望される場合は、日本複写権センター(03-3401-2382)にご
連絡ください。

東大教養学部 編
統計学教室 編 基礎統計学 I
統 計 学 入 門 2800 円

東大教養学部 編
統計学教室 編 基礎統計学 II
人文・社会科学の統計学 2900 円

中 村 隆 英 精 著
新 家 健 泰 人 敬 著
美 添 田 隆 英 精 著
豊 田 隆 英 精 著
統 計 入 門 2400 円

芝 南 純 朝 和 著
南 風 原 朝 和 著
行動科学における統計解析法 3000 円

中 村 隆 英 精 著
新 家 健 泰 人 敬 著
美 添 田 隆 英 精 著
豊 田 隆 英 精 著
経 済 統 計 入 門 第 2 版 2900 円

高 橋 伸 夫 著
経 営 統 計 入 門 3200 円
SASによる組織分析

国 友 直 人 監修
山 本 拓 ・ 編
社会・経済の統計学 4800 円

国 友 直 人 監修
山 本 拓 ・ 編
21世紀の統計科学 I
小 西 貞 則 著
国 友 直 人 编
自然・生物・健康の統計科学 5600 円

国 友 直 人 監修
山 本 拓 ・ 編
21世紀の統計科学 III
北 川 源 四 郎 编
竹 村 彰 通 编
数理・計算の統計科学 5600 円

ここに表示された価格は本体価格です。御購入の
際には消費税が加算されますので御了承下さい。